CS 520 Final: Question 4 - The Big Picture

Diksha Prakash - dp978

December 2019

Abstract

Identifying some of the properties and abilities the chosen bot would need. Identifying at least three issues that would be faced in trying to implement or build such a thing, and possible algorithmic solutions for each.

PsychotherapistBot

1 Abilities

The various abilities expected to be present in the PsychotherapistBot are as follows:

- **The ability to see things.** This would allow the bot to observe human body language and facial expressions and derive the required results from it.

- **The ability to listen**. This would allow the bot to transcribe from speech which would allow further processing of the text. This would also be helpful in dataset generation.

- **Ability to perform sentiment analysis** on sound and text i.e. 'Natural Language Processing'.

- The bot should be **able to converse in multiple languages** i.e. it should be multilingual. This would allow the bot to make a better connection with its patients as humans are more efficient in expressing themselves in their mother tongue. This is also in sync with the whole idea behind psychotherapy a.k.a. 'The Talking Therapy'.

- It should have the **ability to make judgements about the progress**. Hence, it should have a memory to collect data which should be indexed well for processing and ease of access.

- It should be **able to switch from listening to guiding phase** in real time by acting upon the feedback generated from the patients i.e. the response time should be quite low. This would allow to emulate the human-psychotherapist conversation i.e. real-time data processing and response initiation.

- **Ability to acknowledge its threshold and the urgency** to transfer the case to a human-psychotherapist when needed. This would initiate patient welfare and 'service before self' attitude.

2 Properties

The various properties expected to be present in the PsychotherapistBot are as follows:

- It should have a sound and exhaustive knowledge base of the existing disorders and the prescribed actions. It should also have information about the anomalies and how they are to be dealt with. This would ensure correct diagnosis most of the times.

- It should be able to identify different disorders and the severity associated with them.

- It should be able to identify relevant emotions and analyze their source. This approach is an important part of 'Interpersonal Therapy' which is a type of psychotherapy widely used these days.

- In order to allow human intervention in the diagnosis process and provide the most optimal solution, the bot should be able to maintain an extended network of psychotherapists and their specializations and ratings.

- It should generate and maintain detailed notes of each session with every patient to allow analysis and progress tracking.

- It should consider the tone of voice as a critical factor by accessing the framework of 4 dimensions which is used to analyze the tone of voice. The four basic framework used for the analysis are: Funny vs Serious, Formal vs. Casual, Respectful vs. Irreverent and Enthusiastic vs. Matter-of-fact.

- Ability to include real life incidents about people suffering quite similar situation for better connection and impact.

- Human touch to the voice by ensuring honeyed voice with tone variation. This would produce a calming effect and instant connection with the patient.

- Selection of the right jargon and use a persuasive yet unauthoritative speech pattern. Also, the language should be lucid enough for common man.

- Requirement to stay within the legal and ethical boundaries to ensure a personal yet professional experience.

- Maintaining confidentiality by anonymising data during analysis and learning process.

- Taking into account cultural background and avoiding racially biased suggestions. It is very important to understand the ethics and norms of various cultures as what might be okay to someone won't be respectful enough for someone else.

3   Issues

The various issues expected to be present in the PsychotherapistBot are as follows:

- A person with multiple disorders or discomforts. In this case, we need to have the concept of 'Logical Inferences' in the bot to optimize the decision making capability while not leaving any data behind.

- The ability of the bot to converse in multiple languages would require multilingual training of various NLP models. Each model should be trained on an independent and specific corpus as different languages tend to have same words but with distinct meanings.

- Different languages have varied frameworks for analyzing the tone of voices as even the syllables vary. Our dataset should be exhaustive to include all the speech patterns in every language with all the possible tone variations.

- Since the Bot needs to make real time decisions about changing its mode of operation from listening to guiding while interpreting the body language of the humans, key point detection algorithms can be used to identify the body parts and then classify the observed gestures.

- Different gestures can mean different things based on the cultural association. Thus, the Bot should be able to identify the cultural background of the person prior to analyzing the gestures. A possible solution for this is using key point detection algorithms trained on independent datasets for different cultures.

- The real time decision making process is computationally very expensive and thus requires high processing power and highly accessible memory in abundance.

4 Having completed the course, are there any projects you would revisit and improve in light of material you learned later? Be thorough.

1. In Assignment 2 - Minesweeper, obtaining the inference via CNF resulted in $O(2^n)$ complexity which isn't desirable. It could have been solved in polynomial time using Gaussian Elimination or other linear system methods like Gauss Jordan Elimination.

2. In Assignment 4 - Adding a Splash of Color, we had implemented a very basic neural network with 3 hidden layers to produce a fairly accurate colorized image from the grayscale image. The alternatives for the same are:

   - Variation Auto-Encoders: Due to the innate ability of the VAEs to synthesize new data (here, new images) from the given input data (here, input training images), we can expect a better colorization of test images using VAEs.
   - Convolutional Neural Networks: This is an obvious choice as they are able to identify and learn color distributions more accurately than a basic feed-forward network.
   - Datasets like McGill, MIT CVCL, ILSVRC 2015 CLS-LOC, MIR-FLICKR can be used to train our network on a variety of images and colours/hues. In this manner, we can expect the output to be more precise.
     (a) The MIT CVCL Urban and Natural Scene Categories dataset contains several thousand images partitioned into eight categories.
     (b) The McGill Calibrated Colour Image Database contains more than a thousand images of natural scenes organized by categories.
     (c) The ILSVRC 2015 CLS-LOC dataset is the dataset used for the ImageNet challenge in 2015
     (d) The MIRFLICKR dataset comprises 25000 Creative Commons images downloaded from the community photo sharing website Flickr

5 What did my dog dress as for Halloween?

Your dog dressed up as Superwoman on Halloween.

6 Draw a picture of my dog in her Halloween costume.

Figure 1: Professor Cowan's Dog in her Halloween Costume