Deeptanshu Paul
25th February 2020

A Percept-Affect Model of Empathy in Conversational Artificial Intelligence

Abstract

Communication is an important part of all aspects of the human experience especially in its quality. Most businesses spend a large sum of money dedicated to communication with their customers, particularly in the domain of customer service and helplines. To compensate the demand-supply gap of trained personnel most companies are adopting Conversational Artificial Intelligence (AI) in the initial stages of conversation with the customer, and connect to a human only if the issue cannot be resolved through basic Frequently Asked Questions (FAQ). However, present day Conversational AI fall short in one aspect i.e., ignoring emotional cues from the user and saying the same things, regardless of the reaction – not feeling human-like. The hypothesis is that an empathetic approach to Conversational AI, wherein the user's emotional state is given consideration before choosing the utterances that the AI makes, would aid in making conversations seem more human-like. This project is designed with the aim to address this issue using an empathy-centric approach: to **understand**, **process** and **respond** to emotional cues using a conversational interface. The benefits of this approach are that it allows the conversational interface to be better by actively responding to the user's responses instead of being tone-deaf. This translates to reduced time, costs and **first-call resolution** – an important metric for customer satisfaction. The principle this project is based on, is utilizing the power of Deep Learning models to detect emotions in voice data, coupled with sentiment from the explicit intent in the text equivalent of the data, and finally use it to create a robust internal state describing the current emotional state of the conversation. The proposed method has the added advantage of being able to escalate to a human as the situation demands, adding to the effectiveness of the software.

Discussions

This project primarily deals with the emotion recognition and formulating an apt response to the emotion that was recognized. However, the primary motivation, albeit being abstract in nature, is to try to humanize the machine. The term humanizing in this context means being more aware of the factors that affect a human being's communication through the vocabulary they choose to use. The pipeline of the flow of data is not only efficient but maintains a consistency in the way the data is handled with minimal inter-module data transfer. The use of Deep Learning models finds the balance between speed and accuracy as while the training time may be high, the trained model's prediction time is ideal, taking roughly 41 milliseconds for text data. That 41 milliseconds involves the conversion of the text into vectors, predictions of two Deep Learning models, the calculation of Formality and the returning of the output. However, by far the most interesting aspect of this was the Natural Language Generation (NLG) part of the project. Although in the end it was settled that the use of templates was the approach that made the most sense, NLG still is the future scope of the project yet it was highly theoretical in nature. Yet having robust natural language generation that can be leveraged in enterprise applications is a long way off – allowing an algorithm to choose the words may prove a liability. While not documented, several attempts were made to make a generalized NLG module with the intent, entities and actions made into a complete utterance.

The need for emotion recognition is very important in any conversational interface, and the same was confirmed through the course of the project. However, the one unconvincing aspect was that the NLG was template based rather than natural language generation using Deep Learning techniques. While a Variational Autoencoder-based method was tried, it didn't have convincing outcomes, that were satisfactory for the standards set in Industry. The emotion recognition portion of the project is most useful when complemented with an affect mechanism. When the affect mechanism is dictated, it becomes limited to the bounds of the creator's

creativity. Yet this did set a base to build on certainly, the design was made with a much larger goal in mind and the remaining portion would come under future scope of the project. The limitations of using only textual data for the inference of what emotions are being conveyed are still applicable, however. While using context to determine emotions was tried, the unfortunate side effect was error propagation as one erroneous prediction creates a domino effect where the subsequent predictions are affected too. This far outweighed the hypothesized benefit of having context. The generalization ability of Deep Learning based approaches does not mitigate the effects of the lack of context.

The specialization of the sentiment analysis from a simple positive-neutral-negative polarity to emotions definitely helped in making more apt responses - a very sad person was given a different response from a very angry person. The sad person had a more verbose response given to them while the angry person was given the most concise answer available to the system. The ability of the pipeline to be tweaked is, in my opinion, the biggest benefit. If a better approach to emotion analysis is thought of, that model can be very easily plugged in without breaking the system as long as the input and output to the model are kept consistent - both of which are explicitly defined in the form of a state variable.

Conclusions

The conclusion at the end of this project is that emotional awareness in human-computer interfaces (HCI) is a real requirement, especially when the standard to compete with is other humans. If HCI is more pervasive in human lives, the designers and engineers of such interfaces should strive to ensure their applications are more aware of what makes
the human, human. Hence, given that emotions play a huge role in our lives, as brought forth in the introduction, it is a major factor that would play a role in so many decisions that an agent interacting with a human would make. This applies to both physical and software agents. Castellanos et. al.in 2017 [1] try even modelling an emotional state for the agent as well. While emotion is a very abstract concept, when looking at the manifestation of emotion in machines, perhaps modelling the effects of emotion is what should be focused on for the present day understanding of emotions. There is some idea about what the effects of emotions are; and plenty of arguments regarding the same, but different individuals express and are affected by them differently. Hence determining how accurately depicted or modelled emotions are is a fairly grey area. This project, while perhaps isn't an exact model of empathy, is an approach to learned empathic responses, as mentioned by Yalcin and DiPaola in 2018 [2]. The percept-affect mechanism in particular was the most basic and lowest level of Empathy and was decided to be the most apt for the application at hand. More immersive HCI should be deeper in their emotional awareness, and more immersive HCI would be better at detecting emotions as well – facial expressions and body language are very telling of emotions. This project has completed its objective of bringing in more awareness of the user's emotional state and making the chat-based interface more responsive and reactive to what the user says. While the percept-affect mechanism is considered the most basic of cognitive models of Empathy, it is very effective in the domain of text-based dialog communication.

References:

[1] S. Castellanos, L.-F. Rodrıguez, L. A. Castro, and J. Perez, "**Cognitive modulation of appraisal variables in the emotion process of autonomous agents**," *2017 IEEE 16th Int. Conf. Cogn. Informatics Cogn. Comput.*, pp. 102–108, 2017.
[2] Ozge N. Yalcin and S. DiPaola, "**A computational model of empathy for interactive agents**," *Biol. Inspired Cogn. Archit.*, vol. 26, pp. 20–25, 2018.