Emotionally Intelligent Artificially Intelligent Virtual Companion

[1]Mohammad H. R., [2]Yajur S., [3]Keith R., [4]Sunil W.

[1,2,3]Student, [4]Professor, MCT's RGIT with Mumbai & India, [1]hasnain2808@gmail.com,

[2]yajursood25@gmail.com, [3]rebellokeith@outlook.com, [4]sunil.wankhade@mctrgit.ac.in

Abstract The mental health issues of loneliness and depression are two rapidly growing problems in society today. The feeling of being alone is often noted to be an ironic feeling given the almost invasive nature of social networking and mobile platforms today. There is an ever-increasing demand for solutions to help people suffering from these mental health issues. According to government statistics, India needs 13,500 psychiatrists in the country and the nation has only 3,827. Where the country desperately needs 20,250 clinical psychologists, we have only 898. There is also sadly a stigma attached to these problems as Indian society generally looks at people with these problems in a very negative light and outcasts them. AI is a fast-growing phenomenon that has captured the interest of experts from many different fields of study as a way of facilitating complex problem-solving in a way that was not previously possible. This novel solution aims to create artificial intelligence that simulates a lot of the human processes of communication such as establishing long-term relationships with the user, predicting the personality of the user using a model of Mayer's Briggs and generating responses using Openai-gpt2. The response generation model is trained on the Persona-Chat dataset.

Keywords — Artificial Intelligence, Conversational Agents, Deep Learning, Natural Language Generation, Natural Language Processing, Personality Detection.

I. INTRODUCTION

The need to interact with a life that is not our own is natural longing among humankind. As humans, social interaction is what optimized our chance of survival in prehistoric ages and that desire to remain in social groups is what constructed many sociological phenomena like language, behavior, and interaction. Social interaction and the desire to communicate with others is what has made human beings the superior species and it is considered by many academicians as the key to the success of mankind as a species. However, social interaction has served different for mankind in recent decades. While still important for the growth and progress of mankind, social interaction has been promoted on the Maslow hierarchy of needs from basic needs to psychological needs and even self-fulfillment needs for many individuals. People now seek belongingness and esteem from their peers and families. Many even simply seek friendship. A study published by The Economist in September surveyed representative samples of people in three rich countries, namely Japan, UK, and USA, found that 9% of adults in Japan, 22% in the US and 23% in Britain always or often feel lonely, or they felt a lack of companionship, or feel left out or isolated. Loneliness has never really been given much priority in India. According to

an article by Huffington Post, there are many reasons for this—as a much poorer country than, say, Britain, addressing basic physical needs often takes priority over isolation and mental health here. Also, accessing help for even more serious mental health issues continues to be a challenge, and hence, loneliness doesn't get the importance it deserves. The last official count of lonely Indians dates back 14 years. In 2004, the National Sample Survey Office found that 1.23 million men and 3.68 million women were living alone and suffered loneliness. This alone explains the need for solutions to combat such a chronic and devastating problem. This has given rise to the popular subdomains of deep learning and machine learning which have proved immensely useful in a multitude of problems in this field.

II. LITERATURE SURVEY

The Paper Predicting Risk of Suicide Attempts Over Time Through Machine Learning. Clinical Psychological Science [1] proposes a new system that innovatively solves this issue while also resolving other persistent issues in a chatbot interface. Traditional approaches to the prediction of suicide attempts have limited the accuracy and scale of risk detection for these dangerous behaviors.[5][6] The authors of the paper sought to overcome these limitations by applying machine learning to electronic health records

within a large medical database. Participants were 5,167 adult patients with a claim code for self-injury (i.e., ICD- 9, E95x); expert review of records determined that 3,250 patients made a suicide attempt (i.e., cases), and 1,917 patients engaged in self-injury that was non-suicidal, accidental, or non-verifiable (i.e., controls). They developed machine learning algorithms that accurately predicted future suicide attempts (AUC = 0.84, precision = 0.79, recall = 0.95, Brier score = 0.14). Moreover, accuracy improved from 720 days to 7 days before the suicide attempt, and predictor importance shifted across time. These findings represent a step toward accurate and scalable risk detection and provide insight into how suicide attempt risk shifts over time.

People rely on computer agents in their lives, from searching for information to chatting with a bot, to performing everyday tasks. [4] These agent-based systems are our first forays into a world in which machines will assist, teach, counsel, care for, and entertain us. While one could imagine purely rational agents in these roles, this prospect is not attractive for several reasons, which we will outline in this article. The field of affective computing concerns the design and development of computer systems that sense, interpret, adapt, and potentially respond appropriately to human emotions. Here in [2], the authors specifically focus on the design of effective agents and assistants. Emotions play a significant role in our decisions, memory, and well-being. Furthermore, they are critical for facilitating effective communication Key insights include Systems that respond to social and emotional cues can be more engaging and trusted, enabling computers to perform complex tasks in a more socially acceptable manner. Emotionally sentient agents present the exciting potential for large-scale, in-situ experimentation and user experience testing. Large-scale analysis of affective data from everyday contexts is important for improving affective computing systems and helping us learn more about human expression and well-being. It is important for designers to consider the specifications of emotionally aware systems. Learning purely from human-human behavior may not always be the most effective approach and an effective agent may raise users' expectations of competence that the system may not possess.

In [3] the author presents a novel approach for the detection of emotional states from textual data. The considered sentiments are those known as Ekman's basic emotions (Anger, Disgust, Sadness, Happiness, Fear, Surprise). The method is completely unsupervised and it is based on the concept of word embeddings. This technique permits to represent a single word through a vector, giving a mathematical representation of the word's semantic. The focus of the work is to assign the percentage Icf the aforementioned emotions to short sentences. The method

has been tested on a collection of Twitter messages and on the SemEval 2007 news headlines dataset. The entire period is expressed as the mean of the word's vectors that compose the phrase, after preprocessing steps. The sentence representation is finally compared with each emotion's word vector, to find the most representative with respect to the sentence's vector.

III. PROPOSED SYSTEM

By understanding and reviewing the literature and previous attempts at building a system, we propose the following system to address the drawbacks mentioned. As shown in the Fig 1. the objective can be achieved by dividing the system into four parts - 1) The input processor, 2) The personality detection subsystem, 3) The response generation, 4) The output generator

Each input from the user would be first go through basic preprocessing such as tokenization and converted to embeddings in the input processor. The input processor would store the input to list of utterances. If the user sends the input as speech signals, the speech would be converted into text and the other steps remain the same

The personality detection subsystem would consist of a text classifier. The text classifier would classify the text into various personalities using the Mayer's Briggs [add a short description on what Mayer's Briggs is] The tokenized input from the input processor would be fed to a Machine Learning model.

The response generation: This would use the personality detected by the personality detector and prepend it to the list of utterances. This list of utterances would be fed to a Deep Learning based response generation model, that generates a stream of token to be processed by the output generator.

The output processor converts the tokens to English words. These words are displayed in the chat window as text window. Text to Speech is used to convert the text to Speech and the audio is played.
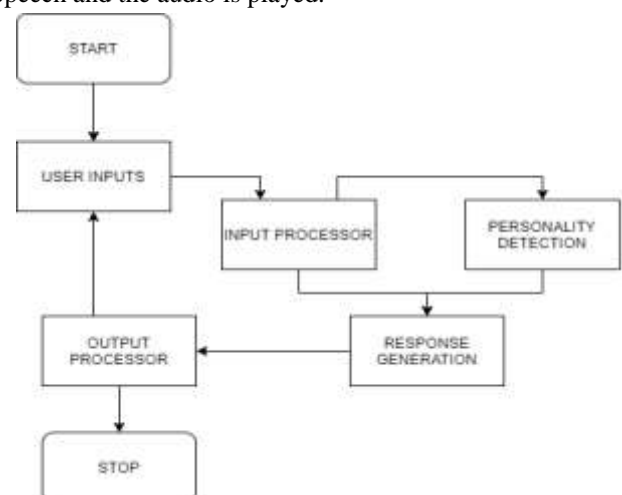

Fig 1. System Flow Diagram

IV. DATASET

A. MBTI dataset

Myers Briggs Type Indicator (MBTI) personality test is a standard in personality assessment. It consists of four parts – Introversion (I) or Extroversion (E), Sensing (S) or Intuitive (N), Thinking (T) or Feeling (F), Judging (J) or Perceiving (P). A person chooses one characteristic from each pair that defines herself/himself properly and then the combination of these characteristics together forms the personality of the test-taker. Therefore, there are a total of 16 different personalities that can be formed using the above metrics and a person can belong to any of the 16 personality types.

The MBTI (Myers Briggs Type Indicator) dataset undertaken for the training of the algorithm consists of 8,600 recorded MBTI personalities of people. The dataset consists of the last 50 posts that the people have posted on social media and against those posts are the MBTI personalities of those people. The people took the MBTI test to determine their personality beforehand and then the dataset was created. This dataset provides the necessary training data required to make a machine learning algorithm that can predict the personality of a person based on the sentence they type. MBTI provides results that can be easily used in a classification algorithm and make the job of creating an algorithm simpler.

B. PersonaChat ConvAI2 Dataset

ConvAI2 Dataset is a set of conversations between randomly chosen crowd workers based on randomly developed personalities that are also randomly assigned to each crowd worker. The conversations are naturally developed with an aim to understand each participant interests and to find a common ground. This task provides an interesting difficulty for the conversational agent to ask and answer questions and to maintain a given personality throughout the conversation. [8]

V. IMPLEMENTATION OF METHODOLOGY

A. Conversation engine

This component performs the tasks of firstly breaking down the user input into base parts that are then passed on to the various other components to further carry out tasks and secondly of deciding the output of the virtual chatbot once it receives information from the other components in the system. The conversation engine consists of an input processor and output constructor. The input processor is the first stage of the conversation engine. It takes user input and assigns tokens to the string of text. Once the input is assigned tokens, the input is merged with the chat history of the user that has been previously collected and stored. The input processor is important as it performs the tasks of

parsing through the text, converting it into useful data that can be easily managed in the other components of the virtual chatbot.

This is the pre-final stage of the conversation engine and the system as a whole as well. The main function of this unit is to construct an output from the other components of the system architecture. It collects the responses from the other components of the chatbot like the knowledge representation manager, depression detection and personality detection modules. It merges these units along with the tokens collected by the chatbot to create a complete input of information for the dialogue generator to construct a response that takes into account multiple aspects of the user's information. It uses OPENAI's GPT2 model to obtain inferences on the output and constructs a reply to the user's input or starts a new conversation depending on the decision taken by inferences generated by GPT2.[7]

B. Personality Detection

This component of the chatbot system architecture deals with determining the personality of the user. The purpose of this is to understand the user's personality type and to also be able to customize the chatbot's personality to better suit the user so that a long-term relationship can be better maintained. This is done with the help of the Myers-Briggs Type Indicator® (MBTI®).

MBTI® is a popular test personality indicator. It consists of 4 qualities which determine the overall personality of the person taking the test. MBTI® is used professionally in companies to know more about their employees and provide them with better work environments or reviews, etc. The four qualities checked in MBTI are – Favorite world (extroversion (E) or introversion (I)), Information (Sensing (S) or Intuition (I)), Decision (Thinking (T) or feeling (F)), Structure (judging (J) or perceiving (P)). Based on the choices of the one taking the test, they fall under one of the 16 MBTI personality types which are all combinations of 4 choices

MBTI® is implemented using the xgboost algorithm in the system. Extreme Gradient Boost Algorithm, or simply XGBoost, is a recently developed machine learning algorithm. As the name specifies, it makes use of three basic algorithms together to make the whole process more efficient, speedy and accurate. Extreme Gradient Boost Algorithm, or simply XGBoost, is a recently developed machine learning algorithm. As the name specifies, it makes use of three basic algorithms together to make the whole process more efficient, speedy and accurate which are explained as follows

Boosting – Models are built in a sequential manner by minimizing the errors from the previous models while increasing (or boosting) the performance of high performing models. By boosting high performing models, the algorithm

ensures that they are repeated more and more while the low-performing models get less frequent. An obvious doubt is that such a process might lead to overfitting but XGBoost algorithm takes care of that as well by regularization.

Gradient Boosting – This is a special case of the boosting algorithm. In this, the errors are minimized using gradient descent algorithm.

Extreme gradient boosting – It powers the gradient boosting algorithm further. Extreme gradient boosting makes use of hardware as well as software optimizations – for example, parallelization, managing internal buffers – to make the algorithm perform faster.

iv.   Using this algorithm, the MBTI® personality type is determined to be one of the 16 personality types. As shown in Fig 3. This information is sent to the output constructor to generate an output that is better able to appeal to the user.

C. Training Process Flow

The Fig 3. depicts the training process flow that has been implemented in order to train the system and its various components. Data is collected from the two respective datasets as described in section 3.3. The MBT dataset is converted to TF-IDF and then trained on an XGBOOST model. The Persona chat data is converted into embeddings and used to predict the personality of of examples of the MBT model. The history of the chat is also taken into consideration and the final model trained is the OPENAI GPT2 model.
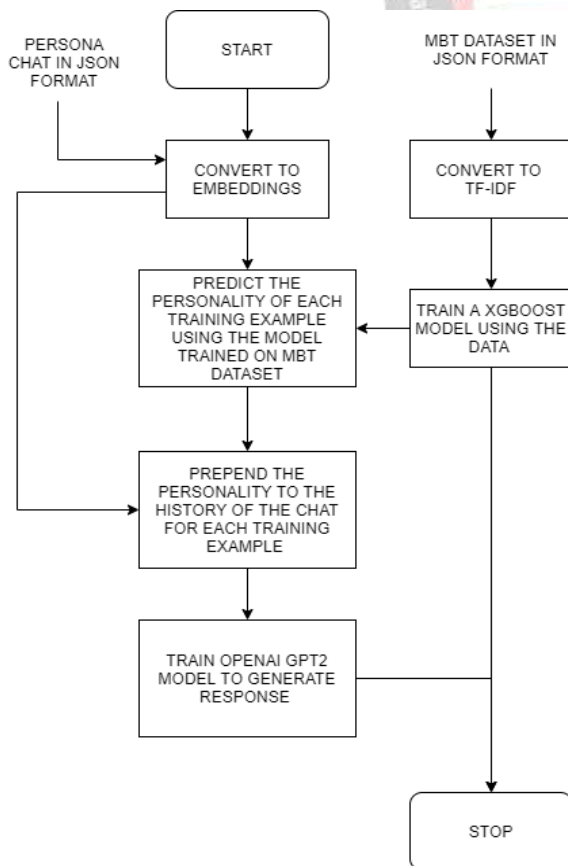
This model uses the data and history to generate an appropriate response the pre-final stage of the conversation engine and the system as a whole as well. The main function of this unit is to construct an output from the other components of the system architecture. It collects the responses from the other components of the chatbot like the knowledge representation manager, depression detection and personality detection modules. It merges these units along with the tokens collected by the chatbot to create a complete input of information for the dialogue.

D. Usage Process Flow

The final operation flow has been depicted in the Fig 4. The system begins with a greeting and it then follows two modes of conversation based on the input mode of the user. If the user chooses to input an utterance the text is parsed and added to the history directly, else if the user chooses to speak the speech is converted into text and sent to the same module as before. The history and the new text is then converted in TF-IDF and also converted into embeddings. The TF-IDF is used to predict the personality of the user and then this personality is added to the history of the user as an example. The embeddings is used with this data to make a prediction for the next utterance and GPT2 is used to generate a response. The response is either displayed as text or it is converted into speech for the user as a response.

Fig 3. Training Process Flow Diagram

Fig 4. Prediction Process Flow Diagram

VI. RESULTS & ANALYSIS

A. Results

The results of the project are reasonably successful. We have presented these results in two ways, namely by presenting the graphs of the training loss learning rate. We also include some metrics such as perplexity. Perplexity is the exponentiation of entropy and a measure of how puzzled the model is. We have also included an excerpt from interaction with bot:

The training and validation loss for the training and validation dataset is 1.585 and 1.501 respectively. The perplexity is 17 .
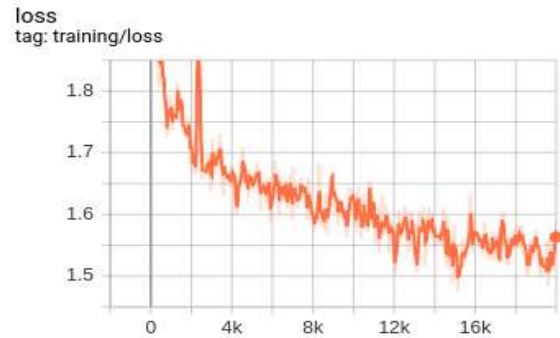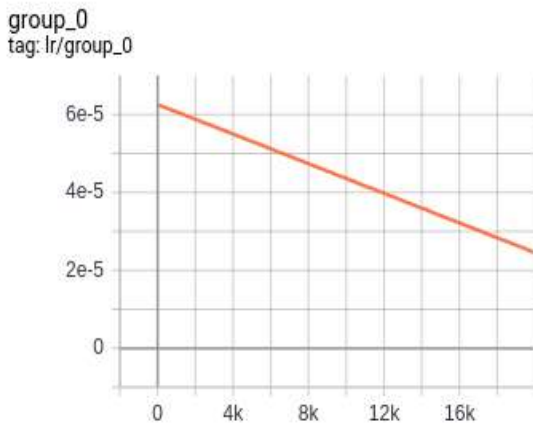

Fig 3. Learning Rate

Following is an excerpt from interaction with bot:
USER: hi
BOT: hello how are you doing?
USER: I am great how are you
BOT: I'm doing well. I'm listening to music.
USER: which music do you like
BOT: I love all music. I love all music.
USER: I like countryside
BOT: I am in the united states
USER: in Texas?
BOT: I'm in Texas
USER: that's great
BOT: I'm also in Canada
USER: how can you be in two places?
BOT: I'm in Canada, how about you?
USER: I am in japan
BOT: that is awesome! my parents are from japan
USER: ohh where in japan?
BOT: I'm from the united states.
USER: your parents?
BOT: my mom and dad are from japan

Fig 3. Training Loss

VII. CONCLUSION

This paper demonstrates that a reasonable degree of emotional intelligence is possible through the system that has been proposed and implemented. We have attempted to build a system that is able to provide Using a combination of models and datasets for the purpose of personality detection, the system is imbued with an enhanced understanding of the user's personality and is able to generate responses that are more emotionally intelligent. While this is a debatable result, it can still to be used as a reference model in further development and research. The inclusion of psychological studies such as the Myers Briggs Type Indicator offers a new prospects of personality detection that is uncommon for typical chatbot systems. This novel inclusion offers a new perspective to chat systems and as demonstrated in this project, does result in an improved conversation with the user which is clearly demonstrated in the results of the project. Through further development and the inclusion of more multi-disciplinary approaches and theories, the scope of growth and development is definitely a possible point of growth that can be considered in the future scope. User's engaging with the system can expect a more emotionally intelligent system that is responsive and articulate enough to meet the scope of this project while remaining in its self-defined limitations.

VIII. FUTURE WORK

There is a lot of scope for future development as mentioned in the conclusion. The inclusion of more human-like features such as mechanical parts or the use of 3D technologies to develop a character to accompany the speech output of the system will enhance its appeal. The systems accuracy can also be improved with further training and the accompaniment of more usage data.

ACKNOWLEDGMENT

REFERENCES

[1] Walsh, Colin & Ribeiro, Jessica & Franklin, Joseph. (2017). Predicting Risk of Suicide Attempts Over Time Through Machine Learning. Clinical Psychological Science. 5. 216770261769156. 10.1177/2167702617691560.

[2] J. McDuff, Daniel & Czerwinski, Mary. (2018). Designing emotionally sentient agents. Communications of the ACM. 61. 74-83. 10.1145/3186591.

[3] Mazzoleni, Mirko & Maroni, Gabriele & Previdi, Fabio. (2017). Unsupervised Learning of Fundamental Emotional States via Word Embeddings. 10.1109/SSCI.2017.8280819.

[4] Zhou, Li & Gao, Jianfeng & Li, Di & Shum, Heung-Yeung. (2018). The Design and Implementation of XiaoIce, an Empathetic Social Chatbot. R. W. Lucky, "Automatic equalization for digital communication," *Bell Syst. Tech. J.*, vol. 44, no. 4, pp. 547–588, Apr. 1965.

[5] Oh, Kyo-Joong & Lee, Dongkun & Ko, Byungsoo & Choi, Ho-Jin. (2017). A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation. 371-375. 10.1109/MDM.2017.64.

[6] Ring, L., & Bickmore, T.W. (2016). An Affectively Aware Virtual Therapist for Depression Counseling.

[7] Wolf, T., Sanh, V., Chaumond, J., & Delangue, C. (2019). TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents. ArXiv, abs/1901.08149.

[8] Dinan, E., Logacheva, V., Malykh, V., Miller, A.H., Shuster, K., Urbanek, J., Kiela, D., Szlam, A., Serban, I., Lowe, R., Prabhumoye, S., Black, A.W., Rudnicky, A.I., Williams, J., Pineau, J., Burtsev, M., & Weston, J. (2020). The Second Conversational Intelligence Challenge (ConvAI2). ArXiv, abs/1902.00098.