



Data Analysis Using NLP to Sense Human Emotions Through Chatbot

Fardin Rizvy Rahat, Md. Shahriar Mim, Aqil Mahmud, Ishraqul Islam,
Mohammed Julfikar Ali Mahbub, Md Humaion Kabir Mehedi^(✉),
and Annajiat Alim Rasel

Department of Computer Science and Engineering, Brac University, 66 Mohakhali,
Dhaka 1212, Bangladesh

{fardin.rizvy.rahat,md.shahriar.mim,mahmud,ishraqul.islam,
mohammed.julfikar.ali.mahbub,humaion.kabir.mehedi}@g.bracu.ac.bd,
annajiat@bracu.ac.bd

Abstract. Years of Technological progress have made machines capable of understanding basic human emotions from their tonality. Written text is also an important method of communication where a wide range of emotions can be expressed. Thus, it is imperative for machines to be also able to understand the emotions being portrayed in the text messages. For this study, an analysis was done on a dataset known as GoEmotions for the potential development of a chatbot. This dataset contains 58 thousand selected Reddit comments collected from various subreddits that use the English language. The results shows that, these were then classified and categorized based on the principal preserved component analysis (PPCA) method into four main emotions; positive, negative, ambiguous, and neutral successfully. And for this task, Natural Language Processing (NLP) pre-training was applied, which was a transformer based machine learning technique.

Keywords: Machine learning · NLP · Emotion recognition · Reddit · Chatbot

1 Introduction

Studying human emotions has become a really hot topic for research in the field of Artificial intelligence. As we know, humans can communicate through a wide range of complex feelings with only a few words, so that is why it has been a goal since the dawn of technology to teach machines how to interpret effects and emotions. Such works can be seen since the 1960s when computer programs were being conceptualized to study communication of natural language between man and machine [1]. Since then, the journey of human speech recognition began through machine learning. Now, through the advancement of technology and with the help of natural language processing we have found new ground to apply artificial intelligence in our daily life. Nowadays, text speech bots are being used

in multiple service-based websites which have eliminated the need for human interactions online. At present, We know some mainstream voice assistants like Alexa, Siri, Google Assistant have become a part of our life [2]. However, these assistants were nothing but command-based operators in our smartphones in their early days [3]. These speech recognizing systems were previously used only to understand certain forms of speech over voice and their capabilities were very limited. Now, they are even able to determine the subtle differences in the tonality of speech and they are gradually becoming closer to understanding user emotions through the way the speech is delivered [4]. This demonstrates great signs of advancements towards the vision of proper machine interpretation of human emotions.

Communication between a man and a computer through emotion and speech recognition has immense benefits. As we know human emotions have a substantial influence on the brain such as the ability of learning, memorizing, problem-solving, perception and different emotional interactions that occur during a conversation affects a person's life and their thinking process. And this is very crucial in modern health care while handling a patient who has stress and depression-related problems through artificial intelligence [5]. Moreover, if it could understand what the user is trying to say as clearly as possible it would really help rehabilitate the patients when they are adapting to their emotional state, thus motivating them and leading the patients to a swift recovery. So, we can say depending upon the state of a patient different approaches can be taken for curing different problems.

Here, in this paper, we studied GoEmotions [6], which is the largest human-annotated dataset containing 58k carefully selected comments extracted from the popular English subreddits on Reddit, where they have been either labeled for 27 emotion categories or as Neutral. Our main goal here is to analyze the dataset in order to use it for building a chatbot that interacts with users' emotions and understands the positive and negative aspects of their delivered speech and tries to promote positivity more in the interactions and remove the negatively perceived emotions from the communications between the chatbot and the user.

The objectives of this paper are mentioned as follows:

- Analysis of all different types of emotion displayed in the dataset.
- Classification, grouping and separation of the data points into binary classes of negative and positive emotions.
- Preparation of the analyzed dataset for chatbot development.

Henceforth, a chatbot can be developed through fulfilling all the objectives, across which human emotions could be understood better in order to help promote more positivity among people.

2 Related Works

With a view to improving the health of Human beings many technological systems and services have been invented. Among them, providing responsive services in a fraction of time, Chatbots developed using Artificial Intelligence is a

very significant invention. Chatbots is a computer program that automatically generates a response in a form of text or speech according to the inputs in texts or speech of the user.

In 1966, J Weizenbaum at MIT invented ELIZA, the world's first Chatbot [1]. ELIZA could run on MIT's MAC time-sharing system only and allowed for some specific types of natural language communication between people and machines. ELIZA evaluated input sentences using decomposition rules that were activated by keywords in the input text. Reassembly rules associated with selected decomposition rules were used to generate the responses. However, ELIZA had some technical problems. For ages, researchers tried to develop more dynamic, robust, and efficient chatbots. Eventually, researchers started to use the Chabot for mental health counseling.

D Lee et al. suggested a novel chatbot system for psychiatric counseling services. They argued that more rational and dynamic sentiment analysis will result in better mental health [7]. They used Natural Language Processing (NLP) to understand the contents of the conversation. They also developed a customized counseling response by taking input from Mental Health Experts. They collected their data from mass media such as scripts of plays and radio shows which contain manually described emotions.

T Kamita et al. [9] proposed the use of smartphones and chatbots to execute a self-guided mental healthcare course. They stated, it is convenient to use smartphones and chatbots for mental healthcare services and works as a motivation for regular and continuous use. They used the SAT method invented by T Munakata which is an interview-style counseling and treatment methodology [8]. SAT is a structured imaging treatment that uses a subjunctive mood to convey people back in time to their ancestors and, finally, to their particles in order to foster the rebuilding of stable attachment links. However, T Kamita et al. had found that using a chatbot may boost a user's motivation and assistance in reduction of stress while also being useful in a self-guided mental health course.

Among all the mental illnesses suicidal thoughts can be very devastating. But thanks to the NLP. In 2016, researcher B L Cook et al. combined the applications of NLP and ML to predict suicidal thoughts and psychiatrist symptoms [13]. Suicidal thoughts and acute psychiatric symptoms were predicted using natural language processing and machine learning among patients recently released from being hospitalized or in emergency hospital settings in Madrid, Spain. At numerous follow-up stages, participants replied to standardized mental and physical health tests. Suicidal thoughts and psychiatric symptoms were the outcome variables of interest (GHQ-12). They contrasted NLP-based algorithms that used an unstructured query with logistic regression estimation techniques that used structured data. Positive Predictive Value, specificity, and sensitivity for NLP-based suicidal ideation models were 0.61, 0.57 and 0.56 compared to 0.73, 0.62, and 0.76 for structured databased models. NLP-based models of elevated mental symptoms had Positive Predictive Value, specificity, and sensitivity of 0.56, 0.60, and 0.59 compared to 0.79, 0.85, and 0.79 in structured models. All in all, they have stated that based merely on replies to a basic general mood inquiry,

NLP-based models were able to yield quite good prediction values. These models have the potential to quickly identify people at risk of suicide or psychiatric distress, and they might be a low-cost screening option in situations where longer structured item surveys are not possible.

Furthermore, in 2021, researcher A Tewari et al. proposed a chatbot and a software named MentalEase and also compared existing chatbots for psychotherapy [14]. According to them, MentalEase is a smartphone application that employs NLP techniques to give not just conversational assistance but also a toolbox of useful features to help people maintain their mental health. Patients with mild anxiety and depression can benefit from integrating mental health evaluation tools into the chatbot interface, as well as a regular treatment. They used fundamental NLP techniques such as word embeddings, sentiment analysis, and models like the Sequence-to-Sequence model and attention mechanism. They compared two chatbots named Wysa and Woebot. Both of them are used for virtual mental therapy. They stated that Wysa, in order to communicate effectively, needs more information and asks a lot of open-ended inquiries. On the other hand, Woebot's communication may appear to be one-sided. It appears to be contrived and does not address many subtle topics. Additionally, they proposed a new chatbot architecture. The suggested system is an Android mobile application that combines the power of Deep Learning and NLP to give not just conversational aid in the form of a chatbot, but also a toolbox of valuable features to help people maintain their mental health. They chose OpenSubtitles Corpus and Cornell Movie Dialog Corpus for the data collection.

The covid-19 pandemic situation causes mental health breakdowns, globally [10]. Researchers felt the urge to develop a chatbot that can do conversations with a concerned user daily. For that reason, Rodrigo et al. [11] proposed a chatbot system that can chat with Students daily to improve their mental health. They used the PERMA model, developed by Martin Seligman [12], to assess a student's well-being and offer coping techniques for maintaining positive mental health and emotions. Their chatbot system communicates with users through a text-based interface.

In our work, we have used a chatbot based on NLP to provide mental counseling determining the user's inputs, since NLP gives the ability to the machines in understanding and process human language so that they may execute repetitive activities automatically and eventually this key functionality of NLP can result in a more robust, accurate, efficient and instantaneous.

3 Dataset

The dataset utilized in this paper is Google's GoEmotions which is obtained from Kaggle [6]. It for the most part centers around conversational information, where feeling is a basic part of the correspondence. Since the Reddit stage offers an enormous, freely accessible content which incorporates direct discussion among users, it is an important asset for emotion investigation. Thus, we assembled GoEmotions utilizing Reddit remarks dated from the beginning of Reddit in

2005 to January 2019, obtained from the subreddits with a minimum of 10k remarks, barring erased data and remarks of other languages.

In order to empower constructing comprehensive agent feeling models, information curation measures were applied to guarantee the dataset does not build up broad, nor feeling explicit, language inclinations. This was especially significant in light of the fact that Reddit is known to have a young demographic predisposition inclining towards the youthful male, which is not indicative of an around the world diverse populace.

Positive		Negative		Ambiguous
admiration 🌟	joy 😄	anger 😡	grief 😞	confusion 😕
amusement 😂	love ❤️	annoyance 😠	nervousness 😰	curiosity 🤔
approval 👍	optimism 🙌	disappointment 😞	remorse 😔	realization 💡
caring 🤗	pride 😏	disapproval 🗨️	sadness 😞	surprise 😲
desire 😍	relief 😌	disgust 🤢		
excitement 🤩		embarrassment 😳		
gratitude 🙏		fear 😨		

Fig. 1. GoEmotions taxonomy including 28 emotion categories

From Fig. 1 we can see GoEmotions taxonomy which consists of 28 emotion categories beforehand. These categories are found after the compilation of all fifty eight thousands selected Reddit comments which was labeled for these 28 categories of emotion. This taxonomy includes a large number of positive, negative and ambiguous emotion categories making it easier to understand tasks that require understanding of emotion such as analysis of customers feedback.

The stage additionally presents a slant towards toxic, hostile language. To address these worries, hurtful remarks were distinguished utilizing predefined terms for hostile grown-up and disgusting substance, and for personality and religion, which was utilized for information separating and veiling. Information was also shifted to diminish irreverence, limit message length, and equilibrium for addressed feelings and opinions. To stay away from over-portrayal of famous subreddits and to guarantee the remarks likewise reflect less dynamic subreddits, the information among subreddit networks was additionally adjusted.

Scientific categorization was made to mutually augment three goals:

- Give the best inclusion of the feelings communicated in Reddit information.
- Give the best inclusion of sorts of passionate articulations.
- Limit the general number of feelings and their cross-over.

Such a scientific classification permits fine-grained understanding that is information-driven and additionally tending to probable information scarcity for particular sentiments.

4 Methodology

In this paper, we will discuss the overall architecture of our model, where the chatbot will feel the emotions from text-based sentences. Here we carried out data pre-processing, feature engineering. The next few paragraphs will demonstrate how the dataset is divided into certain types of emotions such as positive, negative, neutral and ambiguous.

4.1 Dataset Pre-processing

In the stage of preprocessing, we mainly focused on the data cleaning from the raw dataset which excludes id and unnecessary punctuation removal which allows us to get more accurate data.

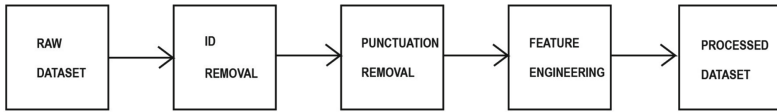


Fig. 2. Steps for data pre-processing

After cleaning the data we carried out feature engineering which in turn allowed us to extract the different features, which is also known as the emotion that will be determined from the sentences via the help of a chatbot. The sequence of steps taken for the pre-processing have been illustrated in Fig. 2.

4.2 Feature Engineering

In this segment of feature extraction from our new processed dataset after pre-processing, we will use this method to extract the different features present.

Table 1. Ambiguous emotions tabulated

confusion	confused-18	why-11	sure-10	what-10	understand 8
curiosity	curious-22	what-18	why-13	how-11	did-10
realization	realize-14	realized-12	realised-7	realization-6	thought-6
surprise	wow-23	surprised-21	wonder-15	shocked-12	omg-11

To carry out the procedure after the process of filtering at least one label amounts to 93% of the raw data. So to carry out training and testing, we split the data into 80% for training and 20% for testing. Thus the final result is based on the test dataset even though we filtered our dataset there were a certain amount of examples that fall under the emotion ambiguous text.

Table 2. Positive emotions tabulated.

admiration	great-24	awesome-32	amazing-30	good-28	beautiful-23
amusement	lol-66	haha-32	funny-27	lmao-21	hilarious-18
approval	agree-24	not-13	don't-12	yes-12	agreed-11
caring	you-12	worry-11	careful-9	stay-9	your-8
desire	wish-29	want-8	wanted-6	could-6	ambitious-4
excitement	excited-21	happy-8	cake-8	wow-8	interesting-7
gratitude	thanks-75	thank-69	for-24	you-18	sharing-17
joy	happy-32	glad-27	enjoy-20	enjoyed-12	fun-12
love	love-76	loved-21	favorite-13	loves-12	like-9
optimism	hope-45	hopefully-19	luck-18	hoping-16	will-8
pride	proud-14	pride-4	accomplishment-4		
relief	glad-5	relieved-4	relieving-4	relief-4	

Table 3. Negative emotions tabulated.

anger	fuck-24	hate-18	fucking-18	angry-11	dare-10
annoyance	annoying-14	stupid-13	fucking-12	shit-10	dumb-9
disappoint-ment	disappointing-11	disappoint-ed-10	bad-9	disappoint-ment-7	unfortunately-7
disapproval	not-16	don't-14	disagree-9	nope-8	doesn't-7
disgust	disgusting-22	awful-14	worst-13	worse-12	weird-9
embarrass-ment	embarrassing-12	shame-11	awkward-10	embarrass-ment-8	embarrassed-7
fear	scared-16	afraid-16	scary-15	terrible-12	terrifying-11
grief	died-6	rip-4			
nervous-ness	nervous-8	worried-8	anxiety-6	anxious-4	worrying-4
remorse	sorry-39	regret-9	apologies-7	apologize-6	guilt-5
sadness	sad-31	sadly-16	sorry-15	painful-10	crying-9

After the completion of training and testing, we found a certain list of words from the list and grouped them into emotions of four categories: ambiguous (Table 1), positive (Table 2), negative (Table 3), and neutral with neutral emotions not used. Hence we will further evaluate the performance of the model on this hierarchy. From the tables above, we are mainly focusing on the positive and negative emotions of our dataset in this paper, that is Table 2 and 3.

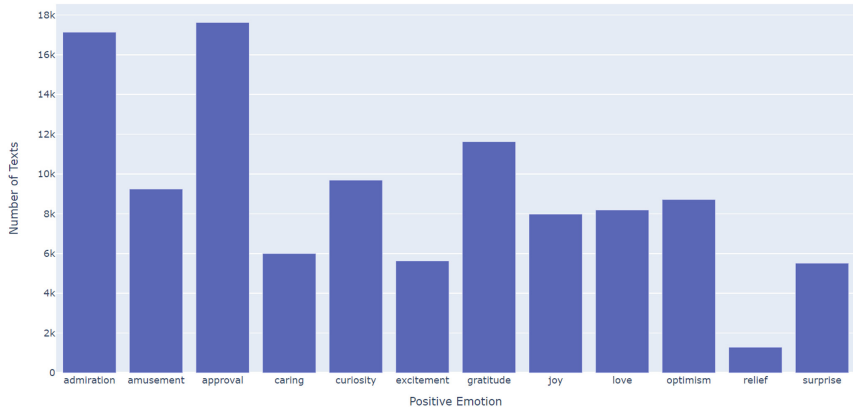


Fig. 3. Positive emotion histogram

In Fig. 3 we can see representation of positive emotions in the form of bar chart which shows a comparison of how many times these emotions were used in texts.

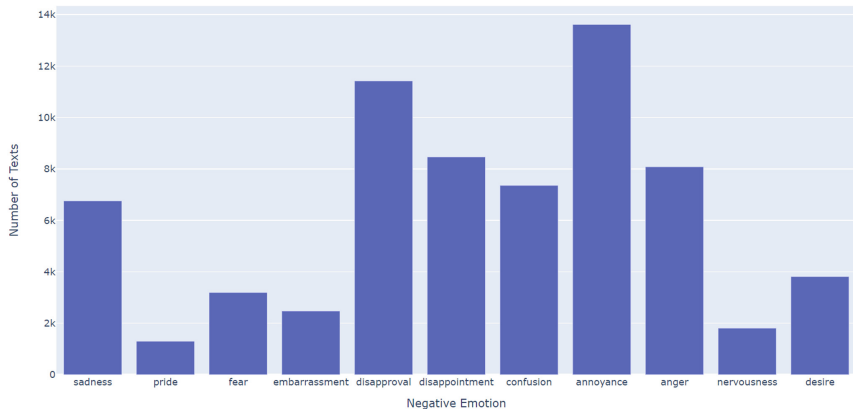


Fig. 4. Negative emotion histogram

The figure shows 12 types of emotions and they are admiration, amusement, approval, caring, curiosity, excitement, gratitude, joy, love, optimism, relief, surprise. Among these emotion approval was used maximum number of times in a text and relief was used minimum number of types.

Moreover, Fig. 4 portrays the negative emotions in the form of bar chart which shows a comparison of how many times these emotions were used in texts. The figure shows 11 types of emotions and they are sadness, pride, fear,

embarrassment, disapproval, disappointment, confusion, annoyance, anger, nervousness, desire. Among these emotion annoyance was used maximum number of times in a text and pride was used minimum number of types.

We do not include neutral emotion since it is not considered part of the semantic space of emotion. Therefore, while carrying out the process we differentiated the types of positive emotion which are demonstrated right below in Fig. 3. Likewise, we differentiated the types of negative emotion which are demonstrated right below in Fig. 4.

5 Result Analysis

In the GoEmotions dataset, emotions are not distributed uniformly. The great frequency of good emotions, in particular, supports our need for a more diverse emotional scientific classification than the authoritative six core feelings. The principle preserved component analysis (PPCA) technique is used to examine two datasets by distinguishing linear combinations of emotion judgment that demonstrate the most elevated joint inconstancy across two arrangements of raters to ensure that the ordered decisions fit the basic data. As a result, it displays emotional elements that are well-understood by all raters. PPCA was already used to understand the main dimension of emotion recognition in video and audio, and it is now utilized to understand the main elements of feeling in text.

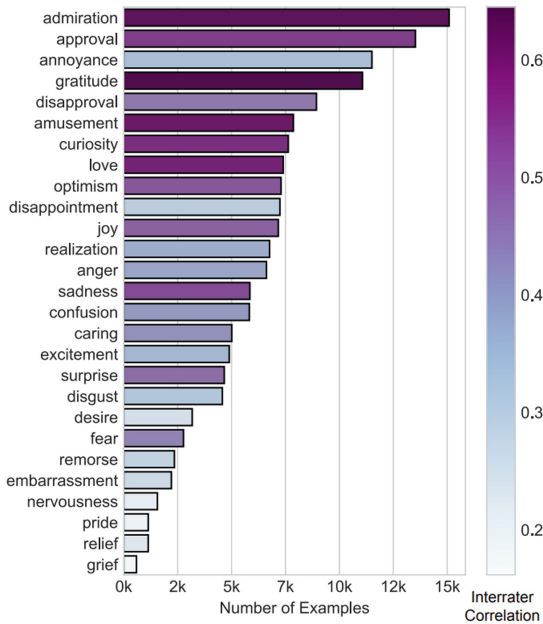


Fig. 5. Chart of frequency of emotions

Every part is observed to be critical (with p-values $< 1.5 \times 10^{-6}$ for all aspects), showing that every feeling catches a remarkable piece of information. This is not paltry, since in prior research on emotion recognition in conversation, just twelve out of thirty components of feeling were deemed critical. The frequency of correlation of the emotions has been shown through the chart in Fig. 5.

We look at the bunching of the characterized feelings dependent on relationships among rater decisions. With this methodology, two feelings will group together when they are as often as possible co-chose by raters.

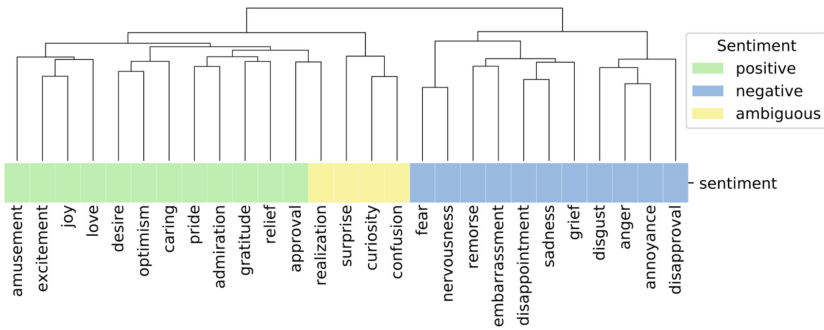


Fig. 6. Grouping of feelings

Despite the lack of a predetermined notion of feeling in our scientific categorization, we see that feelings that are associated in terms of their opinion (positive, negative, and ambiguous) group together, proving the quality and consistency of the evaluations. For instance, in the event that one rater decided “fervor” as a name for a given remark, almost certainly, another rater would pick a connected feeling, for example, “happiness”, rather than, say, “dread”. Maybe shockingly, all uncertain feelings grouped together, and they bunched all the more intimately with positive feelings, shown in Fig. 6.

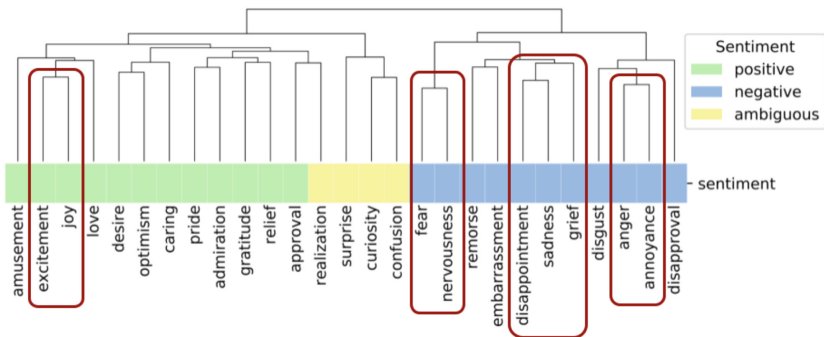


Fig. 7. Connection of sentiments

Essentially, feelings that are connected as far as their competence like joy and excitement, nervousness and fear, annoyance and anger, sadness and grief are likewise firmly related as seen in Fig. 7.

Working out the log odds ratio of all tokens for every feeling class differentiating to any remaining feelings, the lexical correspondences of each feeling were removed. As the log chances are z-scored, all qualities more prominent than three demonstrate a profoundly huge (>3 std) relationship with the comparing feeling. We observe that those feelings that are profoundly fundamentally connected with specific tokens (for example appreciation with “much obliged”, entertainment with “haha”) will generally have the most elevated interrater relationship. Then again, feelings that have to a lesser extent related tokens (for example despondency and anxiety) will more often than not have a low interrater relationship. These outcomes propose specific feelings that are all the more verbally implied and may require additional context to be interpreted.

6 Conclusion

To recapitulate, throughout our research and analyzing the data we reach a conclusion that the chatbot we have attempted to create will help people in the future where they can utilize this chatbot to communicate. For future implementation a fully functional chat bot can be build using the following features along with other features such as- neutral and ambiguous as mentioned earlier. According to GoEmotions dataset, people will input text-based sentences, where the chatbot comes in to detect emotions such as positive and negative emotions which will provide assistance in their daily lives to comprehend human emotions in a sophisticated manner.

References

1. Weizenbaum, J.: Eliza—a computer program for the study of natural language communication between man and machine. *Commun. ACM* **9**(1), 36–45 (1966)
2. Raj, S.: *Building Chatbots with Python: Using Natural Language Processing and Machine Learning*, 1st edn. APRESS, New York (2018)
3. Terzopoulos, G., Satratzemi, M.: Voice assistants and smart speakers in everyday life and in education. *Inform. Educ.* **19**, 473–490 (2020)
4. Jana, A., Patil, H.: The evolution of voice portal and virtual assistants. *Int. J. Emerg. Technol. Innov. Res.* **8**, d556–d564 (2021). ISSN 2349-5162
5. Egger, M., Ley, M., Hanke, S.: Emotion recognition from physiological signal analysis: a review. *Electron. Notes Theor. Comput. Sci.* **343**, 35–55 (2019)
6. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S.: Goemotions: a dataset of fine-grained emotions (2020). <https://doi.org/10.48550/ARXIV.2005.00547>. <https://arxiv.org/abs/2005.00547>
7. Lee, D., Oh, K.-J., Choi, H.-J.: The chatbot feels you-a counseling service using emotional response generation. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 437–440. IEEE (2017)

8. Kamita, T., Ito, T., Matsumoto, A., Munakata, T., Inoue, T.: A chatbot system for mental healthcare based on sat counseling method. *Mob. Inf. Syst.* **2019** (2019)
9. Munakata, T.: Reconstructing life and society with SAT therapy: foundations of the new generation CBT. *Int. J. Struct. Assoc. Tech.* **3**, 35–60 (2009)
10. Cook, B., Progovac, A., Chen, P., Mullin, B., Hou, S., Baca-Garcia, E.: Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. *Comput. Math. Methods Med.* **2016** (2016). <https://doi.org/10.1155/2016/8708434>
11. Tewari, A., Chhabria, A., Khalsa, A.S., Chaudhary, S., Kanal, H.: A survey of mental health chatbots using NLP. Available at SSRN 3833914 (2021)
12. Kontoangelos, K., Economou, M., Papageorgiou, C.: Mental health effects of COVID-19 pandemic: a review of clinical and psychological traits. *Psychiatry Investig.* **17**(6), 491 (2020)
13. Rodrigo, M., et al.: Towards building mental health resilience through storytelling with a chatbot (2021)
14. Seligman, M.: PERMA and the building blocks of well-being. *J. Posit. Psychol.* **13**(4), 333–335 (2018)