ASSESSING SUCCESSES AND OBSTACLES OF SENTIMENT ANALYSIS THE

EFFECT OF NATURAL LANGUAGE PROCESSING ON PSYCHOTHERAPY

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Cory Ooten

October 27, 2022

ADVISORS
Catherine D. Baritaud, Department of Engineering and Society

Communication between humans is a complex and deeply involved process that is constantly evolving over time. A speaker starts with a motivation for communication, they consider what effect they want to occur as a result, and they evaluate the listeners mental state to create an appropriate response. Similarly, a listener must interpret what the speaker means and create a fitting response while also keeping the speaker's mental state in mind. All of this happens many times in the instantaneous moments during communication between humans. A more nuanced part of natural language is conveying emotions during communication. Emotions are a more difficult part of human communication where elements like sarcasm can come into effect. Sarcasm has a multitude of effects on communication, for example allowing someone to say one thing, but imply that they mean the complete opposite. There is an uncountable number of similar hidden meanings that humans put into their communication, often without even meaning to. Richard Socher notes that "tone, sarcasm, idioms, and even emojis" are some of the most complicated parts of human speech today (2018).

As technology has rapidly developed over time, the concept of creating an artificial intelligence that can naturally speak with humans has been a topic often debated. As such, the field of Natural Language Processing (NLP) was formed with the goal of creating an AI that can learn, understand, and produce natural human language (Abdellatif et al., 2022, p. 1). NLP has many uses like translating text from one language to another, programs responding to spoken commands, and summarizing large amounts of text quickly. One of the most commonly seen NLP technologies comes in the form of the virtual assistant that many smart devices include today, like Apple's Siri or Amazon's Alexa. One main concept that Natural Language Processing handles is sentiment analysis. Sentiment analysis is a technique NLP uses to

determine whether data features positive, negative, or neutral emotion. This technique is useful in many applications of NLP, for example if a person is communicating in a very negative fashion, it might be important to select words and phrases that will not provoke or anger the person further.

The technical report this prospectus proposes is concerned with the benefits and roadblocks that sentiment analysis faces as a specific technique. Tightly coupled with the technical report is the STS research paper which focuses on the applications of NLP in the field psychotherapy. Shown in Figure 1 is my expected timeline for completing these components of my thesis:

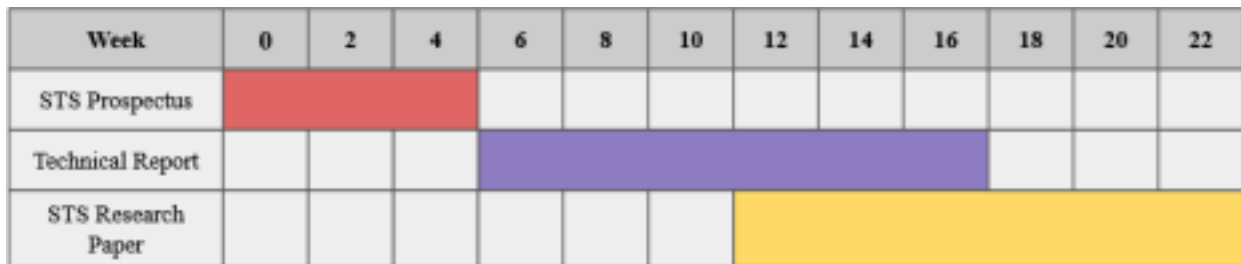| Week	0	2	4	6	8	10	12	14	16	18	20	22
STS Prospectus												
Technical Report												
STS Research Paper												

Figure 1: Gantt chart for thesis components. This figure shows the expected timeline for the completion of the STS Prospectus, Technical Report, and STS Research Paper components of the thesis (Ooten, 2022).

ASSESSING SUCCESSES AND OBSTACLES OF SENTIMENT ANALYSIS

As of August 2022, an average of 6,000 tweets are made per second, equal to 500 million tweets per day (Sayce, 2022). In the modern day, there is an enormous amount of content that expresses and shares opinions about products, media, politics, services, and more. This can be directly attributed to the rise of Twitter and other social media platforms like Facebook and Instagram. Service providers, governments, and private organizations all value what the public is saying about their work, and it is understandable why they look to these

3

platforms and their users for massive amounts feedback (Kaur & Edalati, 2022). However, manually searching through millions of tweets to find ones that mention your company's name is impossibly time-consuming and tedious. Hiring employees to do this can be a waste of resources and they might struggle with accurately analyzing the emotional content included in a tweet. Tweets use all the potentially confusing aspects of human language like sarcasm, idioms, ambiguity, and emojis. The introduction of sentiment analysis helps to ease this pain. Sentiment analysis is an application of NLP with the goal of revealing the emotional state within human speech or text. It is most commonly used to determine the attitude of the speaker toward a topic (Nasukawa & Yi, 2003). Sentiment analysis classifies whether a section of text expresses a positive, negative, or neutral feeling, and is then sub-categorized based on the specific emotion, such as anger or sadness (Mohammad, 2022). Shown in Figure 2 is data from sentiment analysis conducted on Twitter data that references the COVID-19 pandemic:
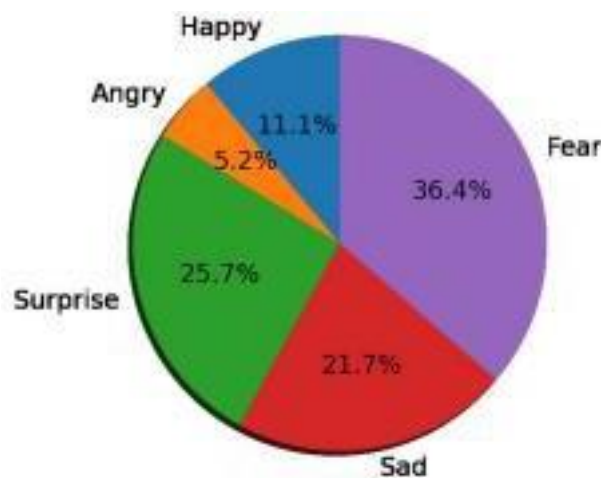

Figure 2: Emotion percentage pie chart. This chart shows the 5 main emotions that the sentiment analysis from Twitter data regarding the COVID-19 pandemic focused on: fear, sadness, surprise, sadness, happiness, and anger (Ramírez-Sáyago, 2020).

However, sentiment analysis is not only used for monitoring social media. It has other uses in fields of market research, voice of the customer analysis, and user services (Zhao et al., 2022). The technical part of the thesis will feature a state-of-the-art report discussing

the benefits and hardships that sentiment analysis faces. The goal of the research is to promote the successes of sentiment analysis and find ways to apply it to new fields, while also discovering where sentiment analysis falls short through the obstacles it encounters. Journal articles related to the field will be the primary sources of information. The main focus of examining the articles will be the areas where implementations of sentiment analysis has fallen short, and where it has met or exceeded expectations. The anticipated outcome of the STS report is a scholarly article to find where projects that use sentiment analysis succeeded and find if their methodologies could apply to designs where sentiment analysis was obstructed.

THE EFFECT OF NATURAL LANGUAGE PROCESSING ON PSYCHOTHERAPY

According to the National Institute of Mental Health (2020), nearly one in five adults in the United States lives with a mental illness. Mental Health America (2022) reports that cases of depression in youths living in the U.S. are growing at a rate nearing 2% annually. One of the main reasons this is significant is that mental illnesses place a heavy burden upon the afflicted. Figure 3 below shows the measured disability-adjusted life-years (DALYs) globally. This is the amount of years a person lives while affected by a mental illness.
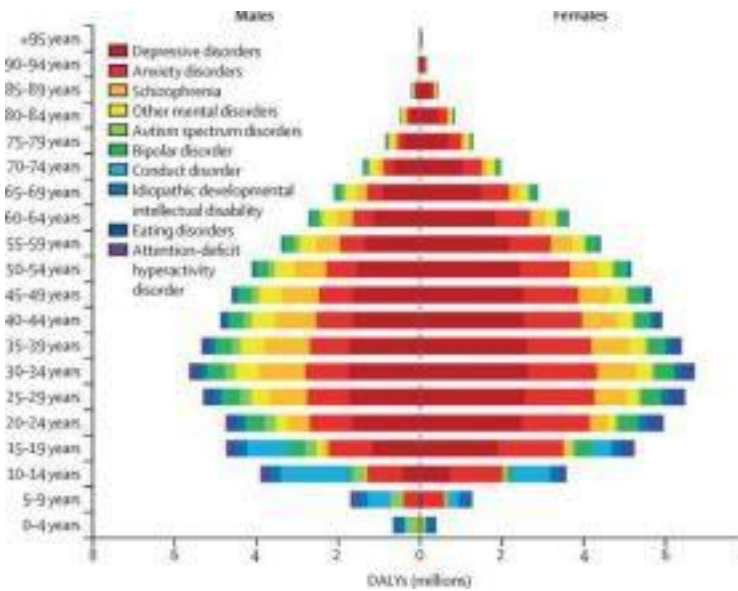
Figure 3: Global DALYs by mental disorder, sex, and age, 2019 (Lancet Psychiatry, 2022).

Mental illnesses can lead to a wide range of symptoms, including consistently feeling sad or dispirited, excessive fears or worries, withdrawal from friends and activities, detachment from reality, major changes in eating habits and more. The fact that so many people live with untreated mental illnesses is not something to ignore.

Despite the effectiveness of counseling and psychotherapy, the number of trained professionals in the field are far outnumbered by the number of affected individuals in the world. In 2019, only 9.5% of U.S. adults received any form of counseling or therapy from a mental health professional (Althoff et al., 2016). Psychotherapy, sometimes referred to as talk-therapy, is a field consisting of a variety of techniques that aim to help a person identify and change troubling emotions, thoughts, and behaviors. This generally involves having a conversation one on-one with a trained profession, or in a one-to-many session with other similar individuals. In a general session lasting 50 minutes, a range of 12,000 to 15,000 words are spoken (Lord et al., 2015). Natural Language Processing provides a mechanism to quickly analyze the transcripts of the conversations to attain a wholistic understanding of patients'

emotions and perspectives. NLP consists of two main components, Natural Language Generation (NLG) and Natural Language Understanding (NLU). NLU specifically refers to reading and comprehending the input of natural human language to the computer and NLG is the computer's output of natural language that a human could understand (Semaan, 2012). NLP has multiple potential uses in psychotherapy, one example being SERMO, a mental health chatbot that is used to help monitor and self-regulate a person's emotions (Denecke et al., 2021).

However, the idea of having an AI monitor your conversations in your most sensitive and vulnerable state might not sound appealing to most people. While recording what you say during psychotherapy and having a machine interpret your true meanings can have real and valuable benefits that psychotherapists might rejoice for, there is also a more hidden complication we must consider. If a patient knows everything they say will be strictly analyzed by an AI, will they say everything they truly mean? Can patients find trust in that everything they say will not be used maliciously later? Can they trust that an AI will not misrepresent their emotions to their therapist? The field of Natural Language Processing requires more research to fully understand the effect it can have on psychotherapy, positive or negative. The research will investigate where NLP is currently being used in psychotherapy by considering how it is being implemented and what positive or negative effects has it had. The research will also search for better ways to use NLP in psychotherapy positively, while minimizing the negative aspects. The objective of the research is to discover the social and ethical concerns with NLP, while also exploring the benefits it provides.

The main approach of the research will follow Pacey's Triangle of Technology (Pacey, 1985). In Pacey's work, he argues that technology is comprised of three main aspects

that work together to create a more complete understanding of the technology, shown in Figure 4. The three aspects Pacey introduces are the technical, cultural, and organizational aspects. Considering NLP, the technical aspect focuses on resources, skill, and techniques requires to produce the technology. The organizational aspect references the institutions responsible for implementing NLP, which is mainly the psychotherapists in this case. And lastly, the cultural aspect is any values, ethical codes, or perceptions that are placed upon the technology by society. All three of these aspects play a major role in the interpretation of Natural Language Processing when used in psychotherapy.

Following the framework of Pacey's Triangle, the research will mostly be concerned with journal articles discussing the applications of NLP as well as studies conducted on real-world implementations of the technology. The anticipated outcome of this scholarly article is to bring to light the uses and benefits of NLP in psychotherapy, while also showing the ethical and social considerations we must have in mind during its production and use.
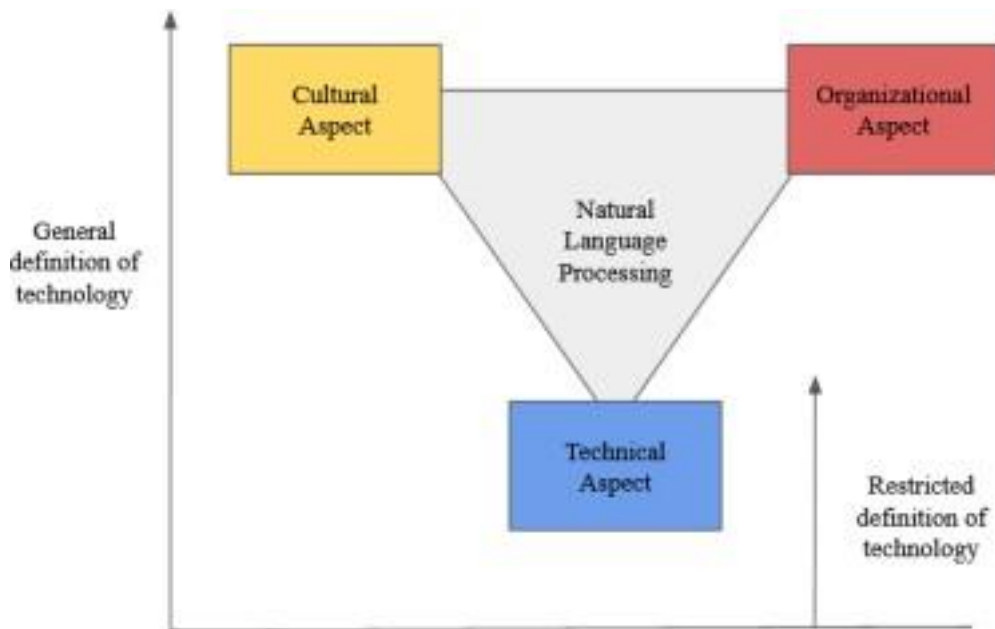

Figure 4: Pacey's Triangle for NLP. This figure shows the balance of the technical, cultural, and organizational aspects surrounding Natural Language Processing (Ooten, 2022).

As technology progresses and evolves, the presence of artificial intelligence that can mimic humans is inevitable. Many instantaneous processes take place the moment that there are even a few simple utterances from humans. Human communication has depth beyond just the words that are spoken. The research proposed in this prospectus aims at understanding the technologies responsible for emulating human communication through artificial intelligence. Specifically, the technique NLP uses, sentiment analysis, is being examined in the state-of-the art technical report for its benefits, drawbacks, and obstacles that it faces in its implementations. The STS research paper is investigating how NLP can be used in psychotherapy and also considers the social and ethical issues that come with it.

REFERENCES

Abdellatif, A., Badran, K., Costa, D. E., & Shihab, E. (2022). A Comparison of Natural Language Understanding Platforms for Chatbots in Software Engineering. *IEEE Transactions on Software Engineering*, 48(8), 3087–3102.

Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health.

Denecke, K., Vaaheesan, S., & Arulnathan, A. (2021). A Mental Health Chatbot for Regulating Emotions (SERMO)—Concept and Usability Test. *IEEE Transactions on Emerging Topics in Computing*, 9(3), 1170–1182.

Kaur, P., & Edalati, M. (2022). Sentiment analysis on electricity twitter posts. *Department of Computer Science Norwegian University of Science and Technology*.

Lancet Psychiatry (2009). *Global DALYs by mental disorder, sex, and age, 2019.* [Figure 3]. Global, regional, and national burden of 12 mental disorders in 204 countries and territories. *Lancet Psychiatry 2022, 1*(9), 137-150.

Lord, S. P., Sheng, E., Imel, Z. E., Baer, J., & Atkins, D. C. (2015). More Than Reflections: Empathy in Motivational Interviewing Includes Language Style Synchrony Between Therapist and Client. Behavior Therapy, 46(3), 296–303.

Mental Illness. (2022). National Institute of Mental Health (NIMH). Retrieved September 22, 2022, from https://www.nimh.nih.gov/health/statistics/mental-illness

Mohammad, S. M. (2022). *Best Practices in the Creation and Use of Emotion Lexicons*.

Nasukawa, T., & Yi, J. (n.d.). Capturing Favorability Using Natural Language Processing. *K*

CAP '03: Proceedings of the 2nd international conference on Knowledge capture. 8. Ooten,

C. (2022). *Gantt chart for thesis components.* [Figure 1]. *Prospectus*

(Unpublished undergraduate thesis). School of Engineering and Applied

Science, University of Virginia. Charlottesville, VA.

Ooten, C. (2022). *Pacey's Triangle for NLP.* [Figure 4]. *Prospectus*

(Unpublished undergraduate thesis). School of Engineering and Applied

Science, University of Virginia. Charlottesville, VA.

Pacey, Arnold. (1985). *The Culture of technology.* Cambridge, MA: The MIT Press. Ramírez-

Sáyago, E. (2020). *Emotion percentage pie chart.* [Figure 2]. Sentiment Analysis from Twitter

Data Regarding the COVID-19 Pandemic. *Tecnológico de Monterre.* Sayce, D. (2022). *The*

Number of tweets per day in 2022. David Sayce.

Socher, R. (2018). *AI's Next Great Challenge: Understanding the Nuances of Language.*

Harvard Business Review.

The State of Mental Health in America. (2022). Mental Health America. Retrieved

September 22, 2022, from https://mhanational.org/issues/state-mental-health-america Zhao,

T., Du, J., Xue, Z., Li, A., & Guan, Z. (2022). *Aspect-Based Sentiment Analysis using Local*

Context Focus Mechanism with DeBERTa. 12.