Machine Learning for Psychopathology Diagnosis

Abdulghani El Masri

Final International University
January 2023
Girne, North Cyprus

Machine Learning for Psychopathology Diagnosis

by

Abdulghani El Masri

A thesis submitted to the Institute of Graduate Studies
in partial fulfillment of the requirements for the Degree of
Master of Science
in
Software Engineering

Final International University
January 2023
Girne, North Cyprus

	FINAL INTERNATIONAL UNIVERSITY **INSTITUTE OF GRADUATE STUDIES**

APPROVAL

Title: *Machine Learning for Psychopathology Diagnosis*

We certify that we approve this thesis submitted in partial fulfillment of requirements for the degree of Master of Science in Software Engineering.

Approval of the Examining Committee:

Prof. Dr. Erden Başar (Advisor) ____________________

Assoc. Prof. Dr. Ece E. Müezzin (Co-advisor) ____________________

Asst. Prof. Dr. Mehtap Köse Ulukök ____________________

Asst. Prof. Dr. Hüseyin Lort ____________________

Approval of the Institute of Graduate Studies:

____

Prof. Dr. Nilgün Sarp
Director

Abdulghani El Masri

abdul.masri@final.edu.tr

ORCID iD: 0000-9911-0099-9009

To my mother and father
for their love and support.

ETHICAL DECLARATION

I, Abdulghani El Masri, hereby, declare that I am the sole author of this thesis and it is my original work. I declare that I have followed ethical standards in collecting and analyzing the data and accurately reported the findings in this thesis. I have also properly credited and cited all the sources included in this work.

Your signature

Abdulghani El Masri

ACKNOWLEDGMENTS

I would like thank to my supervisor Prof. Dr. Erden Başar and co-supervisor Dr. Ece Muezzin for their guidance during the work on this thesis.

ABSTRACT

Mental health has become a rising issue in today's society. Despite being the era where people are most connected to each other, they are still distant. This sense of isolation can invoke many emotions of self depreciation creating now commonly found psychopathological conditions such as depression and anxiety. To cope with this sense of disconnection, people turn to communicating with complete strangers, despite being warned more than a decade ago of the dangers that come with performing such an activity. Additionally, many do not wish to attend therapy sessions due to the feeling of imposition that they have when they visit, as well as the suspicion that the therapy session would be ineffective in quelling whatever they may be feeling.

A solution that psychologists and health experts have come up with is the integration of machine learning to assist them in conversing with and evaluating the patients that they receive. The goal of this thesis is test the effectiveness of such a method and determine how a patient would be able to self-diagnose through a simple text interaction with the chatbot, using a combination of machine learning and data analysis.

There exists many methods to determine the emotions that an individual may be feeling at a given time. From Neural Networks that can read the emotions on a user's face or text to ones that can determine the emotions using simple assignments and calculations to arrive at a conclusion. This thesis will not be implementing an emotional analysis method with Neural Networks due to the Conversational AI already using such a method, and the processing power and

resources required would be too great. As such, the Flask framework has been used to develop the web application, in combination with Meta AI's BlenderBot for the Neural Network based Conversational AI. The greater the model's size, the more processing power required to run it, so the BlenderBot model will be one of the smaller models for the sake of performance. As for the emotional analysis, the NLTK library as well as the Text2Emotion libraries were utilized to create a solution that can measure emotions at a low performance cost.

Keywords: Conversational AI, Neural Network, emotional analysis, psychopathology, mental health

ÖZ

Ruh sağlığı günümüz toplumunda yükselen bir sorun haline geldi. İnsanların birbirine en bağlı olduğu dönem olmasına rağmen hala mesafelidir. Bu izolasyon duygusu, depresyon ve kaygı gibi artık yaygın olarak bulunan psikopatolojik durumlar yaratan birçok kendini değersizleştirme duygusunu çağrıştırabilir.

İnsanlar, bu kopukluk duygusuyla başa çıkmak için, on yıldan uzun bir süre önce böyle bir faaliyeti gerçekleştirmenin getirdiği tehlikeler konusunda uyarılmış olmalarına rağmen, tamamen yabancılarla iletişim kurmaya yöneliyor. Ek olarak, birçok kişi, ziyaret ettiklerinde hissettikleri dayatma duygusu ve terapi seansının hissettikleri her şeyi bastırmada etkisiz kalacağı şüphesi nedeniyle terapi seanslarına katılmak istemezler.

Psikologların ve sağlık uzmanlarının bulduğu bir çözüm, aldıkları hastalarla konuşmalarına ve onları değerlendirmelerine yardımcı olmak için makine öğreniminin entegrasyonudur. Bu tezin amacı, böyle bir yöntemin etkinliğini test etmek ve bir hastanın, makine öğrenimi ve veri analizinin bir kombinasyonunu kullanarak, chatbot ile basit bir metin etkileşimi yoluyla kendi kendine nasıl teşhis koyabileceğini belirlemektir.

Bir bireyin belirli bir zamanda hissedebileceği duyguları belirlemek için birçok yöntem vardır. Bir kullanıcının yüzündeki veya metnindeki duyguları okuyabilen Sinir Ağlarından, bir sonuca varmak için basit atamalar ve hesaplamalar kullanarak duyguları belirleyebilenlere kadar. Bu tez, Etkileşimli Yapay Zeka'nın zaten böyle bir yöntem kullanması nedeniyle Sinir Ağları ile duygusal bir analiz yöntemi uygulamayacaktır ve gerekli işlem gücü ve kaynaklar çok fazla

olacaktır. Bu nedenle Flask çerçevesi, Neural Network tabanlı Conversational AI için Meta AI'nin BlenderBot'u ile birlikte web uygulamasını geliştirmek için kullanıldı. Modelin boyutu ne kadar büyükse, çalıştırmak için o kadar fazla işlem gücü gerekir, bu nedenle BlenderBot modeli, performans açısından daha küçük modellerden biri olacaktır. Duygu analizine gelince, duyguları düşük bir performans maliyetiyle ölçebilen bir çözüm oluşturmak için NLTK kitaplığının yanı sıra Text2Emotion kitaplıklarından yararlanıldı.

Anahtar Kelimeler: Konuşmaya Dayalı Yapay Zeka, Sinir Ağı, duygusal analiz, psikopatoloji, ruh sağlığı

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

- ACT: Acceptance and Commitment Therapy
- AI: Artificial Intelligence
- API: Application Programming Interface
- BERT: Bidirectional Encoder Representation from Transformer
- BLEU: Bilingual Evaluation Understudy
- CBT: Cognitive Behavioral Therapy
- CNN: Convolutional Neural Networks
- CSS: Cascading Style Sheets
- DRAM: Dynamic Random Access Memory
- DSM-5: Diagnostic and Statistical Manual of Mental Disorders 5th Edition
- GAD: Generalized Anxiety Disorders
- GPT: Generative Pre-Training Transformer
- GPU: Graphical Processing Unit
- GRU: Grated Recurrent Units
- HIT: Human Intelligence Tasks
- HTML: Hyoertext Markup Language
- JSON: JavaScript Object Notation
- LSTM: Long-Short Term Memory
- MCBT: Mindfulness-based Cognitive Behavioral Therapy
- MDD: Major Depressive Disorder
- ML: Machine Learning
- NLP: Natural Language Processing
- NLTK: Natural Language Tool Kit

- NLU: Natural Language Understanding
- NN: Neural Networks
- RNN: Recurrent Neural Network
- ROI: Return on Investment
- TPU: Time Processing Unit
- UI: User Interface
- URL: Uniform Resource Locator

CHAPTER 1
INTRODUCTION

1.1 Problem Statement

Most data that is gathered to determine someone's mental health is gathered through a questionnaire. However, the questions that are usually asked are more direct and help identify the symptoms. They also require that a specialist conclude these results to determine a proper diagnosis. An example of a method is the Beck's Depression Inventory, which is a specialized questionnaire that is used to determine the severity of depression that an individual may have (Beck, et al., 1961). An issue that lies with this method, as well as self-questionnaires in general, is that the results that come from it would be prone to bias and human error. Moreover, should the amount of data be sizable, then the processing of said data would be slow as compared to just automating the process, where chances of bias and human error are greatly reduced due to objectivity (Tempelaar et al., 2020). Additionally, there exists methods to determine the emotions of individuals through the use of machine learning. However, these methods require the use of hardware such as cameras and/or microphones in order to gather information on the user based on the user's facial expression, body language, speech patterns, and tone of voice. The accuracy of the method varies based on the implementation, but it is generally highly accurate as a result of using Convolutional Neural Networks (CNN) (Madhavi et al., 2023). Despite the accuracy, it is not an implementation that can not be implemented universally due to the hardware requirements that come with it. Universal implementations would include the use of text, since that would only require the availability of a medium to communicate using unicode. Known implementations are ELIZA, a well-known therapy bot that is more focused on communicating by identifying context through key words, and generating appropriate responses even in the absence of said key words using transformers (Eizenbaum, 1966). Another more modern implementation is Woebot, which aims to act more as a therapist by assisting the users and checking up on them regularly using Cognitive Behavioral Therapy (CBT), all in order to improve the mental state the user is in (Rao, 2018). Both of the aforementioned bots have a set back, and it is that all of their conversations are scripted, none are generative. This means that neither of the bots can identify symptoms of mental illness, they can only follow a list of steps that correspond to the context of their conversation so long as the data to correspond to these conversations is within their dataset. With all of that in mind, how can we develop an AI-based conversational assistant capable of accurately identifying symptoms of depression and anxiety during natural interactions with users and potentially assist individuals seeking support by providing appropriate resources and/or recommendations?

1.2 Purpose of the Study

The main purpose of the study is to determine whether it is possible to analyze someone's general mental state accurately using Natural Language Processing (NLP). Since results in machine learning tend to be more accurate the bigger the dataset, then the use of a modular chat program connected to a conversational AI should prove useful in the extraction of the user's feelings through friendly text. A separate program, that is made for the sole purpose of analyzing the emotion from text, can add to the modularity by making use of the text extracted by the chat program.

1.3 Significance of the Study

There are many potential applications for a conversational AI that has emotion analysis capabilities. Here are just a few examples:

- **Mental Health Support:** An emotion analyzing AI could potentially identify individuals struggling with mental health issues like anxiety or depression, and guide them to resources for support. This technology could also provide additional support to individuals already receiving therapy (Rao, 2018).
- **Customer Service:** Emotional analytics allows businesses to better understand customer needs and tailor their services more effectively. By detecting customer frustration or dissatisfaction, companies can intervene early on and resolve any issues before they escalate (Yam, 2015; Ghazala et al., 2022).
- **Human Resources Management:** HR departments can benefit greatly from emotion recognition systems, improving employee relations, reducing turnover rates, and promoting workplace satisfaction. A happy worker leads to happier customers and increased productivity (Majumder & Mondal, 2021).
- **Gamification:** Imagine designing games tailored to players psychological profiles. Instead of just adjusting difficulty levels or matching similar

interests, developers could actually target feelings themselves through this type of advanced programming techniques. Ultimately it would allow video games to act almost like forms of digital therapy (Rao, 2018).

- **Social Media Intelligence:** Understanding user sentiment can significantly impact advertising ROI. Large social media platforms routinely track this data and monetize according to public mood shifts. A small startup employing an AI assistant would likely face stiff competition from established industry leaders however they could still find niche markets or differentiators offering high accuracy compared to existing methods and thus maintain competitive viability (Ghazala et al., 2022).

These are just a handful of possibilities. As AI becomes increasingly integrated into everyday life, its capacity to interpret and respond to human emotion represents a powerful tool for problem solving and relationship building.

1.4 Research Questions and Hypotheses

A research question would be whether a mental state can accurately be analyzed through text alone, and if so, does growing the chat data make it more accurate when analyzing the user's data? Since machine learning results are said to be more accurate the greater the dataset. As for the hypothesis, through things like vocabulary and negation recognition, along with the use of probability and statistics, an accurate estimate of the user's mental state can be determined assuming the user has been truthful in their chat with the Conversational AI.

1.5 Assumptions

Here are five common assumptions related to conversational AI that can analyze emotions:

1. **Increased Empathy:** Many assume that an AI system capable of identifying and interpreting human emotions will inherently possess greater empathy towards humans than AIs without these abilities. However, while such a feature might enhance an AI's effectiveness in certain areas such as customer service or mental health support, it may be limited to specific situations. Furthermore, artificial intelligence lacks subjective experiences and cannot replicate genuine human empathetic qualities (Coeckelbergh, 2016).

2. **Enhanced Interpersonal Skills:** Another assumption surrounding emotionally aware AI agents is that they would excel at interpersonal interactions. While these systems may indeed improve communication in terms of clarity and sensitivity, actual relationship building remains challenging. Given these AI's lack of authentic physical presence during exchanges, meaningful rapport development beyond professional transactions tends to remain unrealized (Wolff & Stroulia, 2005).

3. **More Human-Like Behavior:** Some believe emotionally intelligent AI will adopt behaviors indistinguishable from real humans, acting naturally even in ambiguous circumstances, expressing nuanced feelings, etc. Yet machines continue operating based upon logic, algorithms and statistical models regardless of contextual variance, thereby limiting the range of human actions available to

them. In most cases, such programs merely process user inputs/responses efficiently (Picard, 1997).

4. **Perfect Sensory Accuracy:** Since AI systems rely heavily on sensors capturing audio-visual cues for analysis, assumptions arise suggesting 100% accuracy when observing human expressions or body language. This overlooks practical challenges posed by lighting, angles, occlusions or other factors impeding smooth perception. Even with refined processing and filtering processes employed by the machine learning network, noise inevitably persists, leading to incorrect conclusions drawn from partial or misleading input signals received. Hence results tend to vary between individual users given the numerous environmental factors possibly interfering with ideal sensor readings (Kaltwang & Maurer, 2006).

5. **Exclusive Affinity Towards Humans:** It's popularly believed conversational AIs trained specifically to recognize emotions in the human species will exclusively focus their attention on people, becoming useless when faced with non-human entities displaying comparable cues (Stomp, 2018).

1.6 Limitations

Several limitations can arise when using machine learning algorithms to recognize emotions from text. Here are a few common ones:

1. **Limited training datasets:** Gathering high-quality annotated datasets can be challenging due to the laborious manual labeling process required. Small sets

may not cover enough diverse scenarios and result types to accurately capture universal emotional themes or nuanced distinctions. Insufficient data volume might hinder performance, particularly when addressing rare categories or specialized domains (Acheampong et al., 2020).

2. **Poor generalization across topics:** Current models primarily learn from preexisting labeled corpora rather than being explicitly designed for specific use cases. Hence, they may struggle with recognizing contextually appropriate states for particular subjects or niche application areas where data is scarce. Adaptation methods are essential for fine-tuning these systems to produce better accuracy (Arjovsky, 2021; Nandwani and Verma, 2021).

3. **Shortcomings handling mixed states:** Humans sometimes convey multiple emotions simultaneously or gradually transition between feelings. Difficulty separating overlapping indicators leads to lower precision in detecting distinct components, possibly leading to confusion about intended expressions. Improved algorithms must handle conflicting cues to enhance their sensitivity to complex psychological states exhibited in subtle linguistic changes (Mehta et al., 2019, Nandwani and Verma, 2018).

4. **Influence of cultural background:** Interpretations of happiness, sadness, fear, anger, disgust, surprise, relief, calmness, contentment, amusement, contempt, boredom, confusion, frustration, irritation, excitement, pleasure, disappointment, interest, etc., vary depending on nationality, ethnic group, language heritage, religion, family environment, societal norms, regional customs, age groups, personality traits, personal history, and many other aspects. Recognition

programs should account for cross-cultural variations in expectations to avoid mislabeling culturally conditioned behaviors or idiomatic expressions. Flexibility in adjusting for contextual factors will help accommodate different worldviews and accents for enhanced cross-domain robustness (Purdy et al., 2019).

5. **Interference from extraneous parameters:** Various components contributing to an utterance's overall meaning can obstruct accurate detection of true emotive subtexts when analyzing chat logs or spoken dialogue recordings containing rich auditory information. Noise interference from background sounds or competitive messages obscuring important words negatively impact reliability. Other sources of corruption include poor audio quality, transmission artifacts, speaker variability, accents blurring phonetic details, dialectal diversity, microphone proximity issues, vocabulary choice, sentence structure, word collocations, tone inflections, rhythm patterns, eloquence, and so forth. Developers must compensate for confounding environmental conditions introduced externally during conversation to minimize negative influences on prediction accuracy (Witek, 2014).

6. **Faulty inference from syntax:** Although parsing text provides valuable clues regarding syntax, punctuation usage, grammar rules, capitalization conventions, orthography standards, spelling consistency, abbreviation preference, syllable length, claw mark count, word frequencies, and other compositional characteristics, too much emphasis on syntactic elements alone risks missing critical semantics hidden within the actual message flow's deeper emotive resonance. Over-reliance on structural cues neglects dynamic aspects of a

conversation such as pauses, silences, hesitation, volume modulations, tremors, and breathing irregularities conveying hidden meanings. Therefore, blending semantic insights harvested from more abstract latent representations generated by deep learning networks combining various input modalities seems necessary for balanced reasoning that captures holistically important emotional undercurrents lying beneath plain surface language constructs (Nandwani and Verma, 2021; Li et al., 2021).

7. **Agnosticism toward unseen emotions:** Machine algorithms face difficulties determining unknown emotional classes since there is no way to guarantee coverage against potential future additions arising from novel circumstances. Even well-documented frameworks may miss identifying emerging states stemming from fresh technologies, evolving customs, or unexpected events transforming communication practices, new forms of self-expression invented online, creative uses of symbolic glyphs, experimental languages mixing existing tongues, innovative memes conveying humorous ideas or references previously unknown, or novel slang spread through viral videos becoming part of daily jargon. Without explicit supervised knowledge incorporating these modern signals, current emotion recognition tools risk remaining ignorant of emerging emotive trends and thus become less effective at comprehensive sentiment analysis. Research focused on proactive monitoring and adaptive evolution to keep up with the continuously evolving nature of human interaction styles could help maintain acceptable levels of perceptiveness. Ultimately, staying ahead of advancing social behavior requires regularly updating machine intelligence

techniques accordingly while also actively soliciting broad public feedback on captured mood impressions for system improvement (Nandwani and Verma, 2018; Mbunge et al., 2022).

1.7 Definition of Key Terminology

- **Machine Learning** a subset of AI that is focused on teaching computers to learn from data and improve from experience (Brown, 2021).

- **Dataset** in machine learning, is a collection of data pieces that can be treated by a computer as a single unit for analytical and prediction purposes (Javaid, 2022).

- **Psychopathology** the study of psychological and behavioral dysfunction occurring in mental illness or in social disorganization (American Psychological Association, 2021).

- **Emotional Analysis** uses a complex system to understand responses made by consumers. Unlike sentiment analysis, where there is generally a positive or negative result, emotional analysis analyzes a much broader spectrum which take into account subtleties within human emotions (Juillion, 2019).

- **Sentiment Analysis** is a system that enables one to understand the general emotions and feelings being experienced by a user, being either positive, negative, or neutral (Juillion, 2019).

- **Negation Recognition** is a process within natural language processing that allows for the detection of words that negate the value of the words that follow them (e.x Not available) (Morante and Blanco, 2021).

- **Natural Language Processing (NLP)** is a branch of artificial intelligence (AI) that enables computers to comprehend, generate, and manipulate human language (Brown, 2021).

CHAPTER 2
LITERATURE REVIEW

2.1 Replika AI’s a Proponent of Pseudo-Companionship

The article by Connor Dewey (2020) presents an overview of the Replika chatbot created by LAION Technologies. Founded by Luan Pinho (2020), Replika generates humanlike interactions by analyzing user input via natural language processing and machine learning algorithms. According to the piece, Replika has gained popularity among teens due to its engaging conversational abilities, which often involve sharing personal stories and asking open-ended questions. However, some critics argue that Replika may serve as a replacement for genuine social connections rather than promoting meaningful relationships. On the other hand, supporters contend that Replika provides users with accessible companionship and serves as a platform for improving mental health by alleviating isolation. Additionally, the article highlights Replika's potential impact in fields like education and customer service. Ultimately, whether Replika represents a useful tool or a concerning trend depends on individual perspectives and intentions when interacting with the technology. Nonetheless, its rapid adoption underscores society's growing reliance upon artificial intelligence and raises ethical considerations regarding our evolving relationship with machines.

In addition to describing Replika, the article also briefly mentions Eugenia Kuyda (2020). After losing her close friend Roman Mazurenko, Kuyda (2020) struggled to come to terms with his loss and eventually decided to create a bot based on his digital footprint. By mining Roman's past communication and social media activity, she fed this data into LAION's neural network to generate a lifelike version of him that people could interact with digitally. Despite initial skepticism, many individuals found comfort talking to the bot since it evoked memories of their departed friend. Kuyda (2020) intended this AI not only to serve as a therapeutic aid but also as an exploration of identity, mortality, and our increasingly intertwined online presence. This concept reflects broader themes in contemporary society around legacy formation, grief management, and the role of technology in shaping how we remember and connect with those who have passed away.

2.2 The Significance of Being Present, The Japanese Man Paid to Do Nothing

An article by Iwata (2021) states that Morimoto Souji (2021) is a Japanese entrepreneur known as "The Man Paid to Do Nothing." He built a successful business model around simply sitting still and doing nothing, which sounds counter-intuitive but has proven effective in attracting clients. This unique approach has garnered considerable attention both domestically and internationally, making him somewhat of a celebrity within Japan's startup ecosystem.

Morimoto (2021) started out as a student studying psychological science at Okayama University. During that time, he became fascinated with the concept of idleness and how it might be applied in a professional setting. He began experimenting with various forms of meditation and mindfulness practices, eventually coming to the conclusion that total stillness was key. In an interview, he said, "[In] idleness, you can feel something hidden deep inside yourself, even if it's just for a moment."

To turn his idea into a viable career path, Morimoto (2021) created a platform called "Achilles," where customers could book sessions with him for private consultations, corporate workshops, or public events. Clients have reported feeling more relaxed and focused after spending time with Morimoto (2021), particularly when they struggle to find inner peace amidst their busy schedules. His popularity grew rapidly, leading to hundreds of requests per week, all thanks to word of mouth alone.

One reason for Morimoto's (2021) success lies in his ability to tap into the growing demand for mental wellness services in Japan. With one of the highest suicide rates in the world and an increasing number of workers facing burnout due to long hours and inflexible job requirements, there is a significant need for solutions that prioritize personal fulfillment over productivity. By offering a unique twist on the familiar concepts of mindfulness and introspection, Morimoto's (2021) Achilles provides a refreshing alternative to mainstream stress reduction methods.

Another factor in Morimoto's (2021) success is the impact of technology on modern society. As people become increasingly reliant on digital devices and information sources, they may find it difficult to disconnect from virtual interactions and engage fully with physical ones. Moreover, the high volume of data transmission and multitasking often lead to sensory overload, causing fatigue and cognitive decline over time. By advocating for a return to simplicity, stillness, and human connection, Morimoto (2021) taps into a collective yearning for balance and authentic communication, especially among younger generations searching for meaning beyond material wealth.

Following the same vein, Morimoto's (2021) philosophy aligns well with the ethos of minimalism, which emphasizes living intentionally by choosing what truly adds value to life and discarding unnecessary distractions. In this context, Morimoto (2021) positions himself not only as a practitioner of mental wellness techniques but also as a guide towards cultivating a lifestyle rooted in mindful choices.

Ultimately, Morimoto's (2021) business model succeeds because it addresses a widespread dilemma arising from the intersection of technological innovation and cultural shifts. While many companies focus on creating new products or services that enhance efficiency or convenience, Morimoto (2021) offers a valuable service in reminding others of the importance of downtime and being present without any obligation or expectation. Through his quiet presence and thoughtful insights, he helps others rediscover their own internal resources and navigate the complexities of contemporary existence with greater clarity.

2.3 Psychopathology Assistant Through the Use of Databases and Cameras

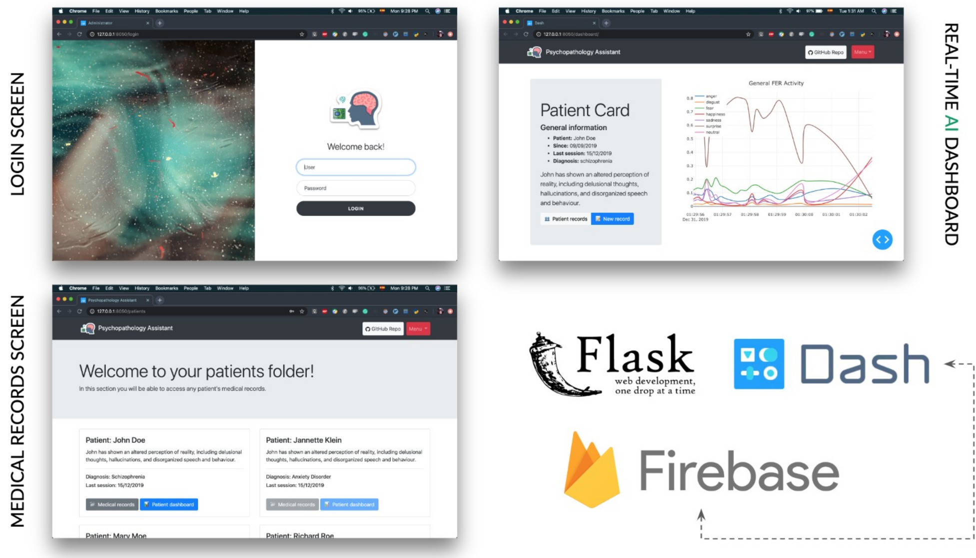


Figure 1: Psychopathology Assistant Web Application

A program was created by Rodolfo Ferro (2019) as an assistant to psychologists so that they may use machine learning to determine what their patient could be feeling during a video call by observing their facial expressions. The program has a login screen for psychologists so that the desired information can be readily available for analysis. The dashboard contains the patients' folders which contain not only their cards, but graphed information of their emotions overtime during the video call. The technology that Ferro (2019) used to develop such a program is a convolutional neural network (CNN) built with the assistance of the Tensorflow library. Ferro (2019) has also incorporated the use of a real-time database called Firebase, as well as the use of a Raspberry Pi Model 3B+ to test the program on a differing hardware.

The author has also incorporated the use of a dataset used in a project known as “Challenges in Representation Learning: Facial Expression Recognition Challenge” by Kaggle.

This dataset has seven categories to that are put to use in the program, which are

- 0 = Angry
- 1 = Disgust
- 2 = Fear
- 3 = Happy
- 4 = Sad
- 5 = Surprise
- 6 = Neutral

According to the author, the program was very satisfactory to members of the jury, for the competition that he took part in, that he was granted victory over other projects.

2.4 Identifying Expression of Emotion Through Text With Deep Learning

A paper by Saima Aman (2007) and Stan Szpakowicz (2007) discusses ways that expression can be identified through the use of text alone. This study proposes a systematic approach for detecting emotional expressions in written content utilizing supervised learning models. Their method involved the usage of six datasets for each of the categories which are *Happiness, Sadness, Anger, Disgust,* and *Surprise.* Each category contains its own seed words to more accurately identify which is which, so for some of the seed words of *Happiness* it would be "happy", "enjoy", and "pleased".

I have to look at life in her perspective, and it would <u>break anyone's heart</u>. (*sadness, high*)
We stayed in a tiny mountain village called Droushia, and these people brought hospitality to <u>incredible</u> new heights. (*surprise, medium*)
But the rest of it came across as a <u>really angry</u>, <u>drunken rant</u>. (*anger, high*)
And I <u>reallllly want</u> to go to Germany – <u>dang</u> terrorists are making flying overseas <u>all scary</u> and <u>annoying</u> and expensive though!! (*mixed emotion, high*)
I <u>hate</u> it when certain people always seem to be better at me in everything they do. (*disgust, low*)
Which, to be honest, was making Brad <u>slightly nervous</u>. (*fear, low*)

Table 1: Annotated Text Examples

This method was dubbed an Annotation Task by the authors, and was usually incorporated with Cohen's kappa to compare the extent of consensus between judges in classifying items into known mutually exclusive categories.

The authors posit that analyzing emoticons in emails and online messages can provide insight into individuals' states of mind, potentially allowing for empathetic responses. To achieve this goal, they compiled annotated datasets consisting of positive, negative, neutral, and non-expressive sentences obtained from the Internet (e.g., online forums) along with corresponding valence scores assigned by domain experts. They then developed several classification models leveraging different linguistic features, feature selection techniques, and ensemble strategies to classify texts as expressive or non-expressive. Evaluating the performance of each setup against various metrics, the authors found that their systems outperformed random chance significantly across multiple tasks. Furthermore, experiments examining the effectiveness of feature combinations, pre-processing steps, and fusion schemes showed improvements compared to standalone models, demonstrating the synergistic benefits of integrating complementary components.

Overall, their results indicate the potential feasibility of developing automatic tools capable of recognizing emotions expressed through language despite limitations posed by noisy input data and contextual variability. Nevertheless, future investigations will be necessary to generalize these findings, account for additional nuances in sentiment analysis, and facilitate broader applications within socially aware computing domains.

2.5 The Creation of an Emotype System for Expressing Emotions With Typeface

The paper proposes a system called "Emotype" for expressing emotions in mobile messenger texting through dynamically changing the weight distribution of characters. The authors argue that existing expressions like emojis, glyphs, and even ASCII art fall short in fully representing complex emotions. They also highlight how modern technology has drastically altered communication norms but hasn't evolved its underlying forms accordingly. Therefore, Emotype addresses three main research goals:

1. Explore the needs and possibilities for expressing emotions through mobile text messaging.
2. Develop an interface for presenting alternative typographic designs when composing a message.
3. Analyze the usability and potential of the prototyped application.

Authored by Choi and Aizawa (2018), the paper studies the usage of typefaces to emphasize on certain emotions from text using a custom mobile messenger. Just like with the Limitations section in the Introduction, they also surmised that emotions are difficult to express without factors like facial expressions, voice tones, or gestures. However that does not mean that expressing emotions is impossible, as the authors have also explained other means of expressing non-verbal signals which involve things like emoticons and capitalization. For example, it is possible to provoke polarization through capitalization by turning “Happy to hear that” to “HAPPY TO HEAR THAT” for the purpose of conveying extreme joy. There is also the letter repetition to extend a syllable like “Sweeeet”, to convey auditory signals which invoke playful impressions. Additionally, emoticons and emojis can help convey an emotion that one might be feeling by completely changing the meaning behind a sentence, so “that’s nice :-)” can convey a positive sentiment where a user may be happy for somebody, while “that’s nice :-(” can convey a negative sentiment that may feel like the user is dissatisfied with another’s action.

Figure 2: Emotype Prototype UI

However, it is through typefaces and emoticons that the Emotype system will express the emotions that a user might be feeling. A prototype of how the user-interface would be presented to the users is seen in Figure 2, where a conversation between two individuals is occurring and a user case for the emoticons and typefaces is being displayed.

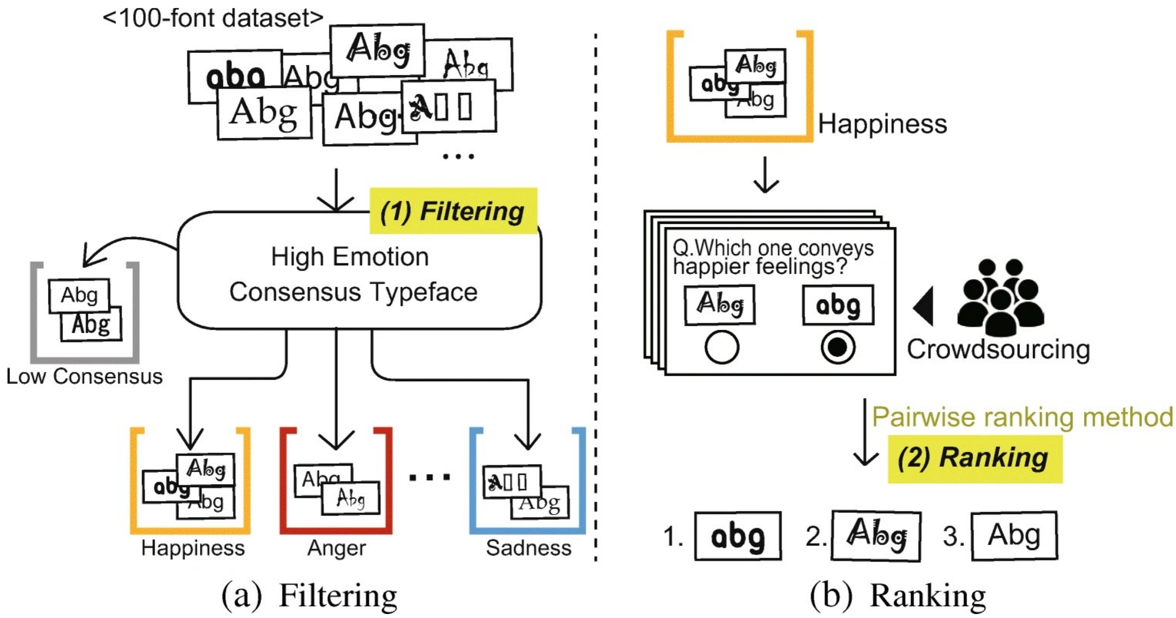


Figure 3: Procedure to Select High Emotion Typefaces

The next figure (Figure 3) shows a basic comparison between static traditional fonts (left side), dynamic variable width fonts (center), and the authors' newly devised variable shape fonts (right). This helps demonstrate the evolutionary step toward what Emotype wishes to achieve.

Slowly, the authors' procedures behind the proof of concept was becoming more and more fleshed out. Figure 4 gives an initial insight into the UI for selecting which mood you wish to convey while typing your normal message. It presents various options including happy, surprised, angry, scared, sad, and love, along with corresponding example images illustrating how the typography responds depending on which choice was selected before composition starts.

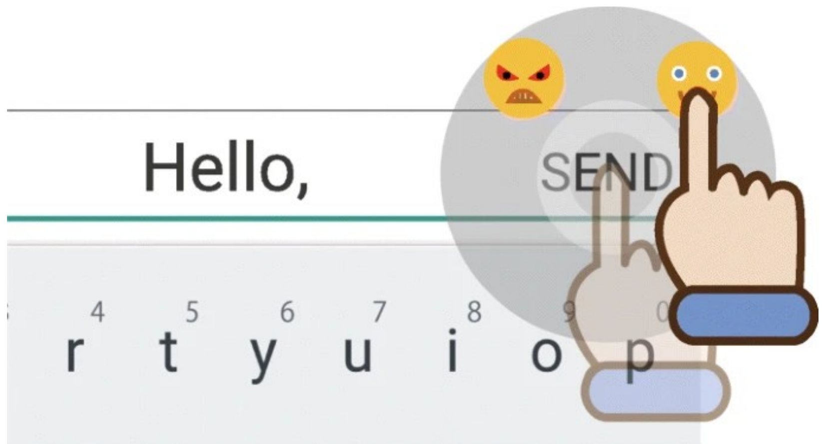


Figure 4: Typeface Changing Interface

Another crucial aspect that has been stated is that Emotype doesn't replace the entirety of what gets sent, just inserts special Unicode codes at designated positions throughout the inputted sentence. These trigger adjustments in letter spacing, width, height, x-height, and weight distributions. This makes it easier for the recipient's device to display the modified appearance once received.

2.6 Dataset Encoding of Text for Deep Learning using Human Intelligence Tasks

A blog by CY Yam (2015) provides a wide range of insight into the methods used to setup a deep learning model that can detect and recognize emotions through text. The author started off by stating that there are six emotional categories that are used widely to describe humans' basic emotions, and they are

Happiness	Disgust
Sadness	Surprise
Fear	Anger

Table 2: Basic Human Emotions

However, these emotions have been re-categorized into four basic emotions which have become

Happiness	Anger/Disgust
Sadness	Fear/Surprise

Table 3: Re-categorized Basic Human Emotions

This provided the challenge of determining how much an emotion can be dependent on a context within a text. This allows the author to bring attention to the lack of labeled emotion databases that allow for active innovation in the topic. Seeing as the two most common databases for emotion recognition, ISEAR and SemEval 2007 databases, do not contain enough data to function as legitimate solutions for emotion detection and recognition. As such, encouragement in this area of study would allow for the development of a more specialized database. However, to develop such a comprehensive database would require vast amounts of data as well as a means to apply a label on the data. To satiate this condition, the use of Mechanical Turks would need to be utilized for Human Intelligence Tasks (HIT). Therefore, the Human Intelligence Tasks would need to be designed with quality control as a priority, hence why there were two designs that have been proposed for such tasks.

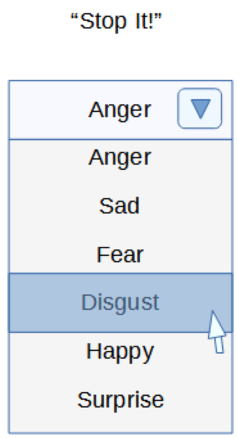


Figure 5: HIT Dropdown Label Selection

In Figure 5, the first design for quality control allows Turkers to select one of the displayed labels for sentences or texts so that they may be better contextualized. Though the design does not allow for a more detailed form of labeling, thus leading to the second design.

Figure 6: Emotion Labeling with Sliders Design

Figure 6 shows the use of the same sentence, although it is easily noticeable that the second design permits a far more detailed labeling of sentences, and that would make it the preferred methodology for purposes related to emotion detection via machine learning.

2.7 Understanding DSM-5 and Diagnosing Depression and Anxiety

Diagnosing and understanding mental illnesses can be complex tasks, especially when dealing with such prevalent conditions like major depressive disorder (MDD). In order to ensure consistency and accuracy in diagnosing MDD, medical professionals rely on guidelines set forth in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). This manual is frequently used by clinicians, insurance companies, and governments around the world to determine the presence, severity, and management of various mental health concerns. As the latest version, the DSM-5 introduced some changes to the diagnostic criteria for MDD compared to its predecessor, DSM-IV. This article by Nancy Schimelpfening (2022), how these updates may influence the diagnostic process and help better understand the nature of depression were discussed.

The current diagnostic criteria for MDD according to DSM-5 are divided into nine distinct symptom clusters. If five or more of these symptoms persist for a minimum of two weeks and cause significant functional impairment in several areas of a person's life, then a diagnosis of MDD may be warranted. These symptom clusters include:

1. Persistent sad, anxious, or irritable mood
2. Decreased interest or pleasure in activities once enjoyed
3. Significant weight loss or gain without effort
4. Insomnia or hypersomnia nearly every day
5. Psychomotor agitation or retardation
6. Fatigue or lack of energy
7. Feelings of worthlessness or hopelessness
8. Impaired thinking, concentration, or decision making
9. Thoughts of death or suicide, suicide plans, or attempted suicide

Once a diagnosis of MDD has been made, the next step involves creating a comprehensive treatment plan. The use of the DSM-5 may influence the approach taken during treatment planning. For instance, if an individual meets fewer than five diagnostic criteria, they might not receive a formal diagnosis of MDD, which would change the focus of their care. Instead, they might benefit from more targeted, less intensive treatments, like cognitive behavioral therapy (CBT), problem solving, or support groups. However, for those meeting full criteria for MDD, a multi-disciplinary approach could be recommended, including antidepressants, psychotherapy, lifestyle adjustments, and other types of therapies or interventions. Additionally, due to specific cultural variations that have become part of the updated DSM-5 criteria, treatment providers should take extra care to consider the context within which patients experience and express their suffering. They can also explore alternative therapeutic approaches to accommodate these nuances.

Another prevalent condition would be Generalized Anxiety Disorder (GAD) which is characterized by persistent and excessive worry about numerous aspects of daily life, accompanied by physical symptoms and functional impairment. To meet the criteria, at least five of the following nine symptoms must persist for at least six months:

Excessive anxiety and worry

1. Muscle tightness
2. Restlessness
3. Irritability
4. Disturbed sleep or insomnia
5. Fatigue
6. Poor concentration or mind going blank
7. Indecisiveness or difficulty taking action
8. Elevated autonomic arousal, e.g., sweating, trembling, palpitations

Some significant changes were made to the categorization and definition of anxiety disorders, including the addition of new disorders such as social anxiety disorder, and changes to existing ones like generalized anxiety disorder. Additionally, some controversial decisions were made, such as merging different subtypes of obsessive-compulsive disorder into one category. Overall, Park and Kim (2020) suggest that further research is needed to improve our understanding of these disorders and refine their diagnoses and treatment. They conclude by suggesting future directions for anxiety research, emphasizing the importance of integrative, multimodal approaches incorporating multiple levels of analysis.

Evidence-based practices play a critical role in treating GAD according to DSM-5. Therapists should draw upon established research findings while considering each client's unique circumstances. Mindfulness-based Cognitive Behavioural Therapy (MCBT), Exposure Therapy, and Applied Relaxation Techniques have demonstrated efficacy in managing GAD symptoms effectively. Clients who do not respond positively to initial intervention choices might benefit from other evidence-based practices like Acceptance and Commitment Therapy (ACT) or other emerging treatments, always keeping the most effective practices at the forefront.

Understanding and properly using the Diagnostic and Statistical Manual of Mental Disorders is crucial for effectively identifying, diagnosing, and treating depression and anxiety. DSM-5 has provided clearer definitions of MDD & GAD and refined their criteria to reflect advancements in the understanding of mental health issues. It allows for greater flexibility and consideration of varied experiences and cultural perspectives, promoting cross-cultural competence. Therefore, it is essential for mental health practitioners and researchers alike to stay informed regarding revisions to this important reference text and apply its principles ethically and responsibly. Overall, being familiar with the latest developments in mental health classification and diagnostics enables a more compassionate, effective, and accurate response to the challenges presented by the prevalence of MDD & GAD (American Psychiatric Association, 2013).

2.8 Unlocking Potential: Using Computer Games to Support Cognitive Behavior Therapy in Children

Cognitive behavioral therapy (CBT) is a popular method used to help children develop coping skills, manage their emotions, and improve overall mental health. One innovative tool that can complement traditional CBT methods is computer games designed to engage young minds while providing valuable lessons in self-regulation and problem solving. This paper explored the benefits of incorporating interactive technologies into CBT sessions with children and review several game examples that can reinforce key therapeutic principles.

Computer games offer numerous advantages as adjunct resources for CBT with kids. They are typically fun, engaging, and can hold the attention of younger clients better than traditional talk therapy alone. Additionally, gaming platforms allow for direct involvement and active participation, promoting experiential learning rather than just theoretical knowledge acquisition. These experiences can be tailored to meet individual needs and interests, enhancing motivation and increasing likelihood of long-term retention. Interactive games can thus serve as effective vehicles for teaching core CBT competencies essential for effective coping in daily life.

Integrating computer games into CBT sessions comes with several benefits. Firstly, it provides opportunities for parallel processing and multiple intelligences activation, meeting the varying learning preferences of participants. Secondly, gameplay allows chances for real-time skill application, where children practice new responses and coping techniques in the safety of the therapeutic setting before generalizing them outside. Finally, collaborating on game solutions fosters social connection, mutual support, and group cohesion among peers undergoing similar challenges. Together, these elements contribute to more enjoyable and meaningful therapy experiences.

Combining cognitive behavioral therapy with interactive computer games presents a unique opportunity to empower children with vital coping skills. Utilizing technology in creative ways not only makes therapy more engaging but also provides access to evidence-based treatment tools. As the digital landscape evolves, therapists should continuously explore novel applications compatible with established CBT frameworks, ensuring they remain relevant in today's tech-driven world. With proper guidance, these dynamic therapy approaches have potential to unlock hidden capabilities within each child, leading to lasting change and brighter futures (Brezinka, 2012).

CHAPTER 3
METHODS

3.1 Research Design

The idea behind this program's design is to consistently extract input information from the user by having them engage with a Conversational AI. Through this method, the amount of information extracted will depend entirely on whether the user decides to stop conversing with the AI, since the AI could continuously take in information and output an appropriate response given the context. Once an acceptable amount of information has been extracted, the program will attempt to determine the user's mental state based on the emotions that have been extracted with the assistance of the Conversational AI. The Conversational AI that will be utilized for this program is Facebook's BlenderBot 2.0, and the method for emotional analysis will be through the Text2Emotion and NLTK Python libraries.

The entirety of the program can be generalized and split into three sections; a section that calls APIs, a section that on the Server-side that acts as a bridge between the APIs and the Front-end section, the final section.

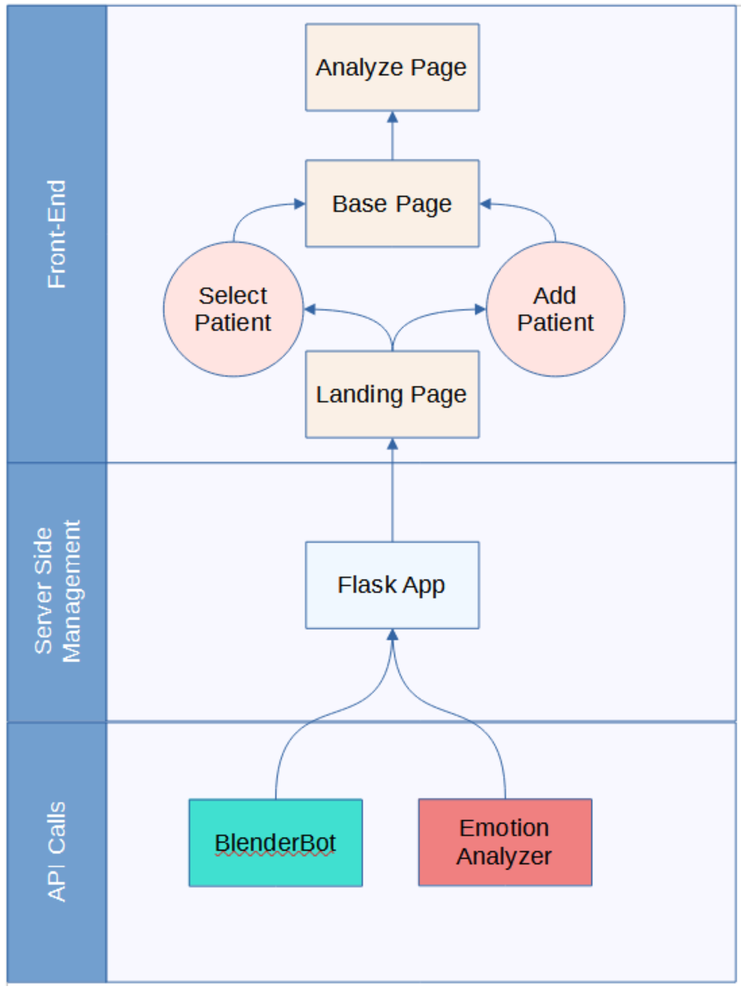


Figure 7: General Structure of the Program

As can be observed in Figure 7, the BlenderBot and Emotional Analyzer are modular programs that are called and managed by the server-side back-end. This server-side back-end connects these APIs to the front-end, as is the standard for web applications. The tech stack for the back-end includes the Flask library, which is a python library for web development that can be used alongside front-end stacks such HTML, CSS, and JavaScript to develop a full-stack application.

3.1.1 Program's Front-end

In accordance to the Front-end section of Figure 7, when this particular web application is run, the user will be met with a landing page where they will be able to chose whether to select an existing patient, or add a new one to start a new session. Whichever option is selected in the Front-end, the request will be a POST request that will be received by the Flask App in the Server-Side Management section. This request will call the URL for the "/base" function, which is a function that will read through all the folders in the *emotion_data* directory, shown in figure 8, and will determine whether to use an existing folder or create a new folder should the new patient's name be non-existent within the directory. The information for these patients will be saved in folders with their name on them, along with excel sheets within these folders marking each session along with emotional data for every input of the conversation.

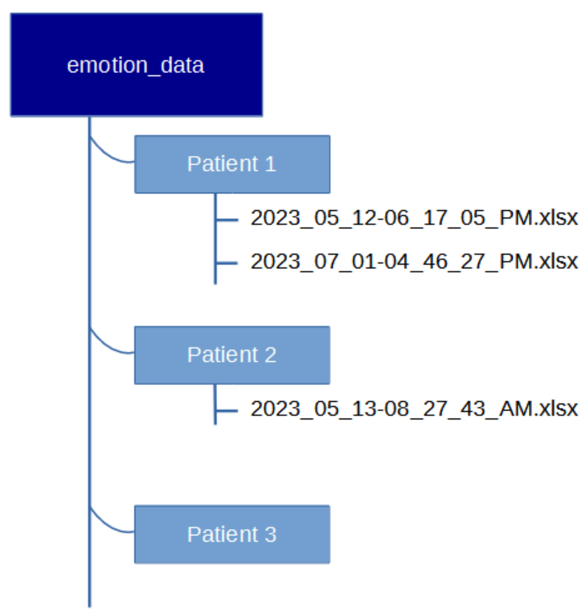


Figure 8: General Patient File Organization

The file structure in Figure 8 can be seen to start with the *emotion_data* folder as the root directory, and every time a new patient is added, the number of *patient* folders increases. Additionally, with every new session that starts for a particular patient, a new excel sheet containing data of the session will be added to the respective patients folder. The excel sheet's naming convention is to indicate the time and date of the session, and the format is as seen in Figure 9.

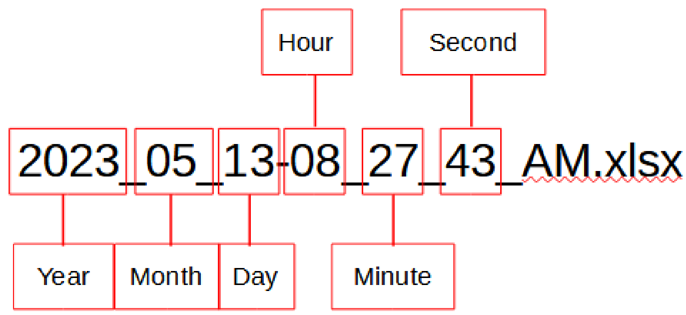


Figure 9: Excel File Naming Convention

It has been implemented in the project directory so that the information can be accessed, modified, and analyzed immediately without the need to run the program. Excel sheets were utilized because it is the most widely used program for spreadsheets and data analyses, so that there would not be any issues with accessing the information of a particular patient.

In the "/base" page, the user will be met with more buttons that allow for the accessing of features within the Flask App, such as the "Analyse" button, the "Merge" button, and the Chatbot's "Send" button, which runs the "/predict" function in the Flask App. If the user were to use the program as intended then wish to analyse a file that they have selected, then they will be redirected to the "/analyse" page along with all the results that should come from having submitted a POST request to the respective function. Otherwise, should the user already have a collection of sessions for a particular patient and decide to merge all of them before analysing the end result, then they would press the “Merge” button to submit a POST request that allows the Flask App to run its “/merge” function and create an excel file that has combined the information of all the files in a particular patient’s folder.

Moving on to the JavaScript functions, these are functions that allow for the uploading of files to be previewed as well as submit a POST request for the “/predict” function via a Chatbox class. Starting off with the excel file management functions there is

- Upload():
    - A function to read the input from the < input type=”file” /> tag to determine whether the selected file is valid. It also calls on the ProcessExcel() function in order to ‘upload’ the excel data onto an HTML table for previewing purposes.
- ProcessExcel():
    - A function that is called on by the Upload() function in order to procedurally generate a table based on the uploaded excel file’s data after said file has been validated by the Upload() function.

As for the Chatbox class, it is meant to be a modular implementation that can call on the Flask App’s “/predict” function to initialize a conversation with the bot, and its functions are

- Constructor():

- As is the standard with constructors, this function initializes the chatbox by selecting all the related UI elements, as well as emptying its messages and setting its active state to false.

- Display():
  - The following function is the chatbox UI itself, it has event listeners for every feature that the chatbox uses, like becoming visible at the press of a button by setting the state of the chatbox to true, or activating the function that sends the POST requests for the chatbot to converse.
- toggleState(chatbox):
  - A function that sets the state of the chatbox, so if it is currently in the HTML page and its state is set to false, then it will remove the chatbox from the page and vice versa.
- onSendButton(chatbox):
  - This function allows the chatbox to submit POST requests by taking in the text from the < input type=”text” /> tag and converting it into a JSON format in order for Flasks “/predict” function to read it as input and return a JSON output of its own, while only calling the updateChatText(chatbox) function when it receives a response from the bot.
- updateChatText(chatbox):

    - The UI of the chatbox is updated using this function. It determines whether the sender is the bot or the user and dynamically appends the DIV into the chatbox messages class.

The "/analyse" page also contains JavaScript functions. However the functions within that page use a JavaScript library called "ChartJS" to plot the data that is received from the Flask App using "Jinja2".

3.1.2 Server Side Management

The Server Side Management section is where the program takes place. Using the Flask Python library, a popular web framework that used specifically for web development, it can create GET and POST requests as well as use "Jinja2" to communicate with the Front-end while also allowing for the flexibility of Python to be used in the Back-end. To be able to use the functions for the API calls, it has imported the scripts as an extension of its own, thus allowing for the usage of functions in other scripts within in its own. The functions that constitute this section and allow it to manage all of the program's features are

- index_get():
    - This function is to initialize the information in the *emotion_data* folder by reading them and appending them into an array so that when the landing page is rendered, the information is transferred to the Jinja protocols in the "Landing Page" file.
- base():

- Once the POST request is submitted by one of the Front-end's "Landing Page" buttons, this function initializes the by clearing the data array indicating a new session is starting with a particular patient whose information gets requested. If the request receives an empty text input, then it is redirected to the "Landing Page", other wise it will create a new excel file using a date-time format for the naming convention and add it to the respective patient's folder.

- predict():
  - As stated in the previous section, this function receives a JSON request from the chatbox in the Front-end, and with the assistance of the emotional analyzer from the API calls, it creates a JSON format of the analyzed emotions' results as related to the text input. The JSON results are then appended into an array which is then inserted into its respective generated excel file, with the assistance of the Pandas data management library, overwriting previous inserted information. This means that so long as the session is active, the information within the excel file is reflective of the information within the array. Once all of those steps are completed, it receives a reply from the bot by using the imported Blenderbot's "get_response(text, history)" function, then converting the response into a JSON format to submit it onto the aforementioned Chatbox class.
- analyse():

- This function will read the file that has been submitted by the Front-end's < input type="file" /> after a POST request is submitted. The file's content will then be analyzed using the functions from the Emotional Analyzer to determine whether the results in the file show the user having signs of Depression or Anxiety, then inserting the results into a JSON format. Afterwards, the information is compiled by finding the mean of each column and saving the result as a JSON format. Once the "Analyse" page is rendered, the information on Depression and Anxiety as well as the compiled emotion information will be submitted for use in "Jinja2", to determine what mental state the patient is in, and "ChartJS", to plot the data of the compiled information.

- merge():

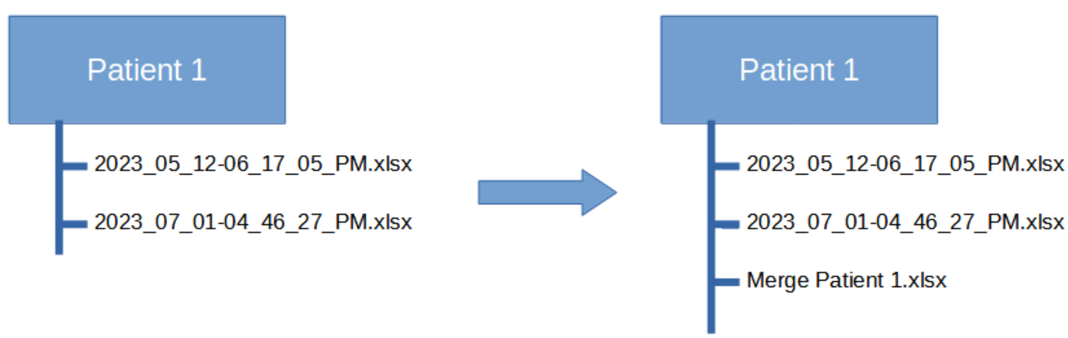


Figure 10: Merge File Arrangement Result

- Once the "Merge" button in the "Landing Page" is pressed, this function receives a POST request and proceed to put all of the respective patient's excel files into a file list. The file list is then converted into a Pandas dataframe using a for loop and Pandas' read_excel() function, with each new dataframe being appended into an array. The array is the concatenated in order to merge all the information together, and finally a new excel file is created with the name "Merge {Patient Name}.xlsx" inside the patient's folder.

3.1.3 Functionality of API Calls

As for the API Calls, their functionality has been displayed in previous sections, showing that they are responsible for information gathering as well as interacting with the user of the program. The way that the program registers the information through the chatbot is presented in Figure 11.

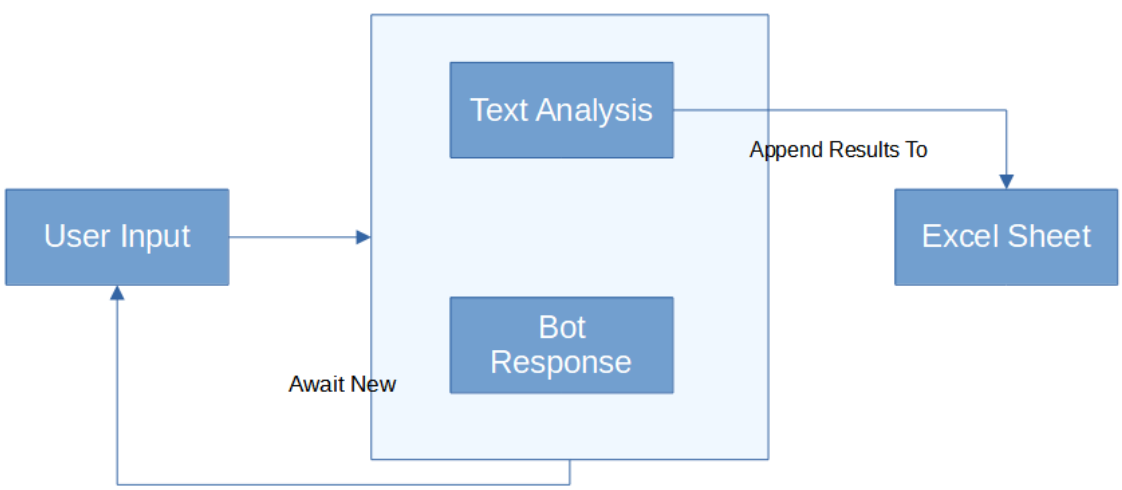


Figure 11: Chatbot Information Gathering Pipeline

As stated prior, the pipeline in the figure is how the information is gathered from the chatbot and put to use by the program. The user would input a text, which would prompt the program to respond to the text via the BlenderBot while analyzing the text and appending the results into the excel file that would generate for that particular session.

The main functionalities in the API Call section are, as observed, the Conversational AI, BlenderBot, and the Emotional Analyzer. To start off, an explanation will first be provided on the functions of the Emotional Analyzer class and the Analyzer class. It uses a combination of NLTK library's Vader Sentiment Analyzer, as well as the Lexicon-based Text2Emotion library to extract emotional information from the patient.

- sentiment_analysis(text):
    - A simple function within the Emotional Analyzer class to provide polarity scores on a given text, through the use of the Vader NLTK library. It returns polarity scores for Positive, Negative, Neutral, and Compound scores.
- emotional_analysis(text):
    - Yet another simple function within the Emotional Analyzer class in which the Text2Emotion library is being utilized to return values for Happy, Sad, Angry, Fear, and Surprise emotions through a lexicon-based approach.

Moving onto the Analyzer class, it contains functions for calculating results for Depression and Anxiety based on the excel files.

- depression_analysis(directory):
    - A function to calculate depression by comparing the difference in results of the mean of the Happy column and the Sad column, after reading an excel file from the directory variable.
- anxiety_analysis(directory):
    - A function that calculates the mean of the Fear column and the Compound column to determine whether someone has anxiety. If the Compound is negative and the Fear is greater than a certain threshold, then the patient has anxiety.

Moving on to the functions of BlenderBot, its functionalities were called using the transformers library of the HuggingFace platform.

- take_last_tokens(inputs, note_history, history):
    - Filters the last 128 tokens so that they are more readable when it is inserted into history.
- add_note_to_history(note, note_history):
    - Add notes to history, for bot short term memory.
- get_response(message, history):
    - Function to receive the bot's response after having it filtered and added to history by means of the aforementioned functions.

3.1.4 Methods of Extracting Emotion Through Text

There exists two known methods for the classification of emotion from text, these methods are a Lexicon-based Method and a Deep Learning Method. Unlike its Deep Learning counterpart, The Lexicon-based Method is heavily reliant on the vocabulary used in a sentence when conducting the analysis (Mohammad, 2011; Gupta, 2020).

Input: I was encouraged by my teacher

Output: ['joy', 'trust']

Figure 12: Lexicon-based Method I/O

Figure 12 is an example of what a Lexicon-based approach would look like. As stated prior, the approach relies on the vocabulary present in the input, and the output would display an array of emotions related to the word from a dataset. There are numerous benefits that come with taking a lexicon-based approach to text emotion analysis. These advantages include:

- **Simplicity:** Lexicon based methods tend to be simple to implement. Given the right dataset, building accurate word categories may be all you need to develop an effective sentiment analysis tool. Since your approach will be grounded directly on words, rather than concepts, no extra steps must go into identifying or defining entities like products, people, companies etc (Mohammad, 2011).
- **Flexibility:** Text documents can vary considerably depending on subject matter and genre, yet with carefully selected dictionaries, lexicons can account for differences between types of texts without any changes in design. Word classes can incorporate different sets of vocabulary across a range of topics and styles. Even within one document or sentence, different senses and connotations of particular words become applicable.

Without explicitly recognizing the entities themselves, your tool could take those factors into account, depending on the quality and breadth of its reference material (Mohammad, 2011).

- **Cost Effectiveness:** Because they typically use freely available resources online, most lexicon based methods are accessible cost-wise too. While larger datasets behind modern machine learning algorithms can sometimes run well into millions of dollars or pounds, good dictionaries are usually free, meaning anyone can start exploring emotion detection without breaking the bank first. And given that your classification is based chiefly on exact matches, storage requirements won't stretch terribly far beyond what's required for your source materials anyway. Therefore you should have few overheads to worry about. It's entirely feasible to build something functional with a minimum budget (Mohammad, 2011).

While lexicon based systems have a lot going for them, there are some drawbacks too. The following points address some issues associated with this type of text emotion analysis approach:

- **Rigidity:** A major downside to relying on preset lists of keywords lies in the inflexibility it imposes upon your application's capabilities. Any shifts in people's attitudes towards emotions (for instance as a result of cultural changes over time), demands for novel kinds of analyses or fresh emotion labels, all represent significant challenges to overcome using only fixed word banks. There exists little room for maneuvering once established unless you opt for retraining altogether. Whereas machine learning models have sufficient capacity to manage unexpected inputs or accommodate evolving criteria since their parameters adjust during runtime (Mohammad, 2011; Mbunge et al., 2022).

- **Shortcoming:** An important issue connected with employing predefined emotional lexicons comes with the possibility that some relevant terms remain absent or omitted due to forgetting by catalogers or failure to detect such nuanced distinctions made by speakers/writers. Depending on exactly which list you work off, crucial subtleties between similar feelings, connotations, or interpersonal situations could escape detection because they weren't previously known or covered in collected dictionaries. Especially in English, countless synonyms plus context-sensitive usages imply that attempting to identify emotions by means of words alone poses steep difficulties to begin with. Machine learning solutions benefit from increased adaptability to unknown scenarios owing to the fact that users can fine tune their training data in response to emerging deficiencies (Mohammed, 2019; Li et al., 2021).
- **Scarceness:** Pertinent resources offering extensive coverage on diverse languages might prove hard or expensive acquire or license. Apart from availability and cost factors, cultural differences also make it difficult to apply a universal set of lexical rules for all regions where emotional expressions vary significantly due to regional idiosyncrasies shaping how people perceive and express feelings. Some countries demonstrate

varying degrees of reservedness versus openness, which manifest in distinct speech patterns and content emphasis. These variances further challenge the usefulness of predefined word lists purportedly capturing every nuance imaginable because certain terms deemed appropriate in one linguistic tradition might feel strange or even offensive elsewhere. Customizing classifiers accordingly calls for fine-tuned understanding of local preferences, customs, dialects, colloquialisms, humor styles, historical backgrounds, religious norms, political beliefs, social networks, interpersonal interactions, group dynamics, family structures, education levels, generational attitudes, economic status markers, gender roles, sexual orientations, physical surroundings, weather conditions, time zones, holidays, festivals, popular entertainment genres, literary works, culinary specialties, et al. (Acheampong et al., 2020).

As for the Deep Learning approach, it makes use of the following Common methods

- Long Short Term Memory (LSTM)
- Bidirectional Encoder Representation and Transformers (BERT)
- Universal Sentence Encoder

As their names imply, all of the aforementioned Deep Learning Methods rely on encoding to classify the emotions of a given sentence. One of the encoding techniques works as follows

Happy	0
Sad	1
Angry	2
Fear	3

Table 4: Encoding Emotions Table

The emotions are first given numerical representations via a table which is going to be used as a reference point for the dataset that contains the training data, so such a dataset will look like

Phrase	Emotion
"There is something outside, and it does not sound human."	3
"I am not quite feeling it today, the death of that one character really got to me."	1
"There is no much for me to be joyful about!"	0
"I can't believe that they would do this, and betray the fans of the franchise like that!"	2

Table 5: Encoded Emotions Usage Dataset

As shown in Table 5, the dataset consists of phrases and emotion encodings to give these phrases emotions that may have been felt when the phrases were stated. The encoding method employed depends heavily on the type of neural network used. Most commonly, for textual input, embedding layers are included in the first few layers to encode words into numerical vectors representing relationships between them. Embedding allows capturing semantic nuances beyond basic bag-of-words representations that ignore word order information. By transforming discrete tokens into continuous representations, it effectively improves neural network capacity and efficiency for handling text data.

There are several advantages to leveraging Deep Learning techniques for emotional analysis compared to traditional rule-based systems or shallow machine learning methods:

1. **Higher accuracy:** The ability to automatically learn complex representations from raw data enables Deep Learning models to achieve higher accuracy in detecting subtle differences between emotional states than purely rule-based or feature engineering-driven strategies. Deep Learning's strong capacity to discover hidden patterns helps optimize predictions by identifying previously unknown relationships. In essence, Deep Learning allows for much deeper mining of available information, potentially uncovering intricate dependencies unnoticed before (Santosh et al., 2022).
2. **Scalability:** Being trained once on large amounts of data empowers Deep Learning architectures to generate generalized representations capable of tackling various contexts with minimal additional adjustments required for each specific domain. Once designed and trained, these versatile frameworks may serve multiple purposes, making them highly scalable while reducing total effort dedicated solely to adaptation processes. Their transfer learning capabilities extend applicability to similar problems without needing extensive model retuning or extensive labeled samples, enabling easy integration into new domains.

Additionally, fine-grained tuning of hyperparameters often leads to relatively small changes throughout many applications, simplifying maintenance costs (Santosh et al., 2022; Biswas, 2021).

3. **Adaptiveness:** Automatic emotional intensity quantification poses unique challenges due to language variation caused by evolving slang usage, emerging platforms, changing opinions, temporal events, and other dynamic aspects that complicate static rulebook-driven detection. However, modern Deep Learning algorithms like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), Transformer series, or BERT employ advanced training regimens encouraging their models to remain adaptable and adjust according to fresh input they haven't seen during initial preparation phases. As the underlying principles for generating outputs come closer to resembling a "black box" whose inputs produce desired effects with little knowledge about internal behavior, deep learning has faced questions regarding transparency and interpretability. Although these issues have been addressed partially through visualizations like heat maps or gradient-based sensitivity analyses, the exact reasoning driving decisions made by these models remains difficult to ascertain (Biswas, 2021; Justus et al., 2018).

While deep learning offers several advantages over traditional rule-based systems or shallow machine learning methods for textual emotional analysis, it also faces some drawbacks:

1. **Data-intensive:** Deep learning requires copious amounts of high-quality annotated training data, which can be a challenge to collect, verify, and maintain. Creating and curating suitable datasets for training deep learning models is time-consuming, requiring substantial resources, including labor, storage space, compute power, and specialists skilled in data collection, annotation, and management techniques. Limited label coverage or imbalanced distributions might also impact model performances if not handled properly (Santosh et al., 2022).
2. **Model complexity:** Building effective deep neural networks typically involves trial and error, experimentation with different architectures, configurations, hyperparameter settings, regularization techniques, activation functions, optimizers, and so forth, until satisfactory results emerge. Given the combinatorial nature of these elements, finding optimal configurations is an iterative process prone to local minimum pitfalls and difficulties reproducing precise architectures. Furthermore, extremely complex models may become computationally prohibitive, either in terms of latency constraints or resource limitations for online deployment. Some architectures might even yield diminishing returns after achieving sufficient expressivity, negatively affecting further increases in depth and width without noticeable improvements (Santosh et al., 2022).
3. **Black box transparency concerns:** Despite recent developments aimed at boosting explainability, deep learning still struggles to expose how predictions arise from mathematical operations performed upon inputs.

This obstacle arises when attempting to decipher internal logic governing conclusions reached or producing feature representations, obscurity surrounding learned patterns could hamper interpretation attempts. Lacking clear justifications for critical decisions made by emotion detection software raises ethical considerations if applied to sensitive scenarios involving healthcare, legal proceedings, hiring processes, or social welfare assessment. In those cases, unexplainable results could create trust barriers between people and artificial intelligence unless explicit explanations clarify how algorithms reach specific outcomes (Rudin and Radin, 2019).

4. **Hardware and runtime costs:** Scaling up deep learning models tends to increase computational budgets needed for efficient execution because more parameters imply longer inference times and memory requirements. Larger architectures necessitate higher levels of parallelism, frequently favoring GPUs or TPUs for accelerated matrix computations, although increasing precision needs may push performance bottlenecks to DRAM speeds. Running these high-end GPUs, especially for large-scale deployments or frequently updating/refreshing models, imposes significant power consumption (Biswas, 2021).

With all of these factors considered, the Lexicon-based approach has been selected for this project, namely due to the low hardware usage and wide range of applicability. Therefore, an appropriate library to perform the analysis task will need to be selected, and the Python libraries available are NRCLex and Text2Emotion.

3.1.5 Differences Between NRCLex and Text2Emotion Libraries

The NRCLex (Bailey, 2019) library is focused primarily on building Natural Language Understanding (NLU) applications powered by large language models, whereas the Text2Emotion (Gupta, 2020) library serves as a lightweight extension allowing Python developers to quickly build text-to-emoji APIs integrated with machine learning frameworks and pipelines.

While the former provides functionalities to train and deploy custom NLU models using Hugging Face, PyTorch, and JAX, the latter functions as a higher order component wrapping together several deep learning libraries, including TensorFlow, PyTorch, Keras, Scikit-learn, and XGBoost among others. Additionally, Text2Emotion focuses more on real-time, efficient application deployment using webhooks, endpoints, middleware, API gateways, message brokers, containers, Kubernetes, serverless computing, cloud services, edge microservices, event driven architectures, and streaming analytics platforms, enabling business continuity, compliance, and scalability in production environments.

Furthermore, Text2Emotion supports multi-threaded parallelism using thread pools, asynchronous execution, coroutines, green threads, gevent, locks, queues, futures, promises, signals, clocks, timers, notifications, logging, monitoring, tracing, profiling, testing, feedback loops, error handling, backpressure, circuit breaking, bulkhead isolation, timeouts, retries, fallbacks, and load balancing in shared memory and distributed computing scenarios.

Finally, Text2Emotion extends capabilities beyond emoticons, offering integration with chatbot avatars, 3D graphics, animations, AR/VR visualizations, and multimedia content generation suited for next-generation virtual world experiences. Both packages share similar objectives yet differ widely in terms of scope and complexity, with NRCLex emphasizing large scale language modeling pipelines suitable for academic experiments while Text2Emotion targets practical enterprise applications requiring robustness, scalability, reliability, security, and maintainability.

3.1.6 Blenderbot 2.0 as the Conversational AI of Choice

"Recipes for Building Open Domain Chatbots: BlenderBot 2.0", as detailed in the article by Weston and Shuster (2021), presents a novel methodology for creating open domain conversational agents capable of generating coherent responses without relying exclusively on predefined templates or strict patterns. Unlike traditional chatbots restricted to specific topics, BlenderBot combines state-of-the-art language understanding components with generative models conditioned on human interaction histories through reinforcement learning principles guided by user preferences. By designing flexible architectures adaptable to diverse use cases, this study proposes a modular framework facilitating future development and innovation in conversational AI.

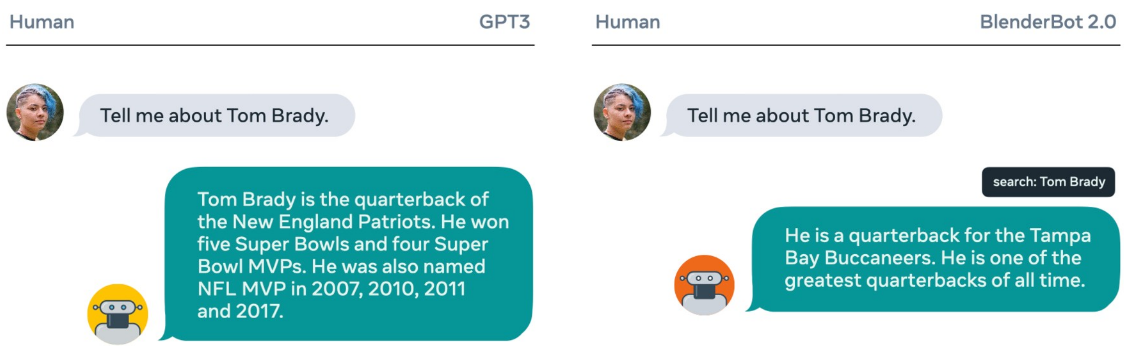


Figure 13: GPT3 and BlenderBot Response Comparison

To achieve their objective, the authors introduce several key improvements upon earlier versions of BlenderBot centered around dialogue policy refinement and response diversification. Firstly, they employ human evaluation measures focusing on fluency and relevancy to optimize reward functions measuring desired properties for each stage of conversation management. Secondly, instead of merely choosing one answer candidate per round, BlenderBot now generates multiple candidates during beam search operations maximizing likelihood under respective distributions, reducing repetition artifacts inherent to greedy sampling regimes. Thirdly, inspired by modern human communication conventions, the system introduces a small fixed probability for abstaining or declining certain turns when uncertain about appropriate replies, avoiding forced sequences leading to sub-optimal outcomes.

Lastly, additional technical contributions include integrating continuous prompt engineering into training schedules, utilizing improved model pre-training paradigms, and adopting effective stopping criteria minimizing time wasted on converged models unlikely to yield better overall results.

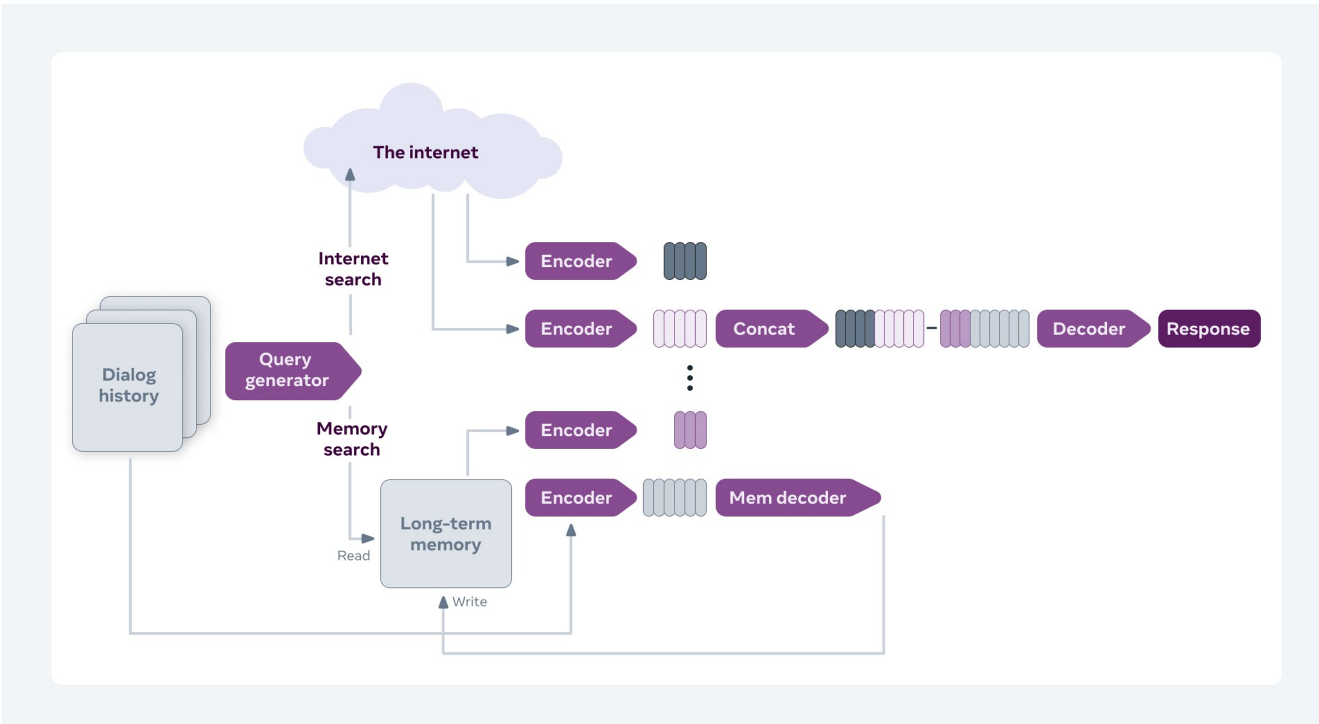


Figure 14: BlenderBot 2.0 Model Architecture

The paper discusses experiments comparing BlenderBot against established baseline chatbots demonstrating clear advantages in terms of human judgments and automatic metrics such as BLEU scores and distinctiveness ratios. By providing detailed guidelines alongside publicly released codebases, This study addresses an important issue faced by modern chatbots, which often suffer from lack of context awareness, inadequate generalization ability, and over-reliance on explicit training data. BlenderBot attempts to overcome these shortcomings by leveraging sophisticated pretrained Transformer language models combined with minimal task-specific fine-tuning. While impressive progress has been made throughout, there still exist opportunities for further research on improving the accuracy and interpretability of conversational systems in challenging domains demanding high fidelity reasoning, common sense, and domain expertise. For instance, potential future work could investigate techniques for integrating knowledge graphs, semantic parsing, commonsense reasoning, or even zero-shot transfer across disjoint domains via meta-learning approaches. Another fruitful direction might entail developing explainable chatbots exposing their decision mechanisms, uncertainty estimates, or confidence rankings to users willing to understand how generated answers were derived. Such endeavors would require significant advancements in natural language processing and machine learning, but ultimately promise even greater societal impact in assisting human interactions, augmenting intelligence, and fostering collective wisdom at unprecedented scales.

Ultimately, "BlenderBot 2.0" represents an exciting breakthrough pushing forward the boundaries of conversational artificial intelligence, opening up promising new paths toward smarter, more intuitive, and communicatively agile digital helpmates serving humans effectively in their daily lives.

3.2 Population and Sampling

The collection of data samples will be based on each individual user, since as discussed above, the main aim is to attempt a pseudo-companionship while analyzing their emotions through their text inputs. Each text input, along with emotions analyzed for it, will be considered a sample of a local user.

The emotions that are being analyzed are:

- Happy
- Sad

- Surprise
- Fear
- Angry

Along with sentiments:

- Positive
- Neutral
- Negative

Where each of these will contain probabilistic information regarding each text input provided by the user. All of the information will be presented in the following format

Text	Happy	Surprise	Sad	Angry	Fear	Negative	Neutral	Positive	Compound
<User inserted text>	0.0	0.0	0.0	0.0	1.0	-0.85	0.15	0.0	-0.7
<User inserted text>	0.0	0.0	0.5	0.0	0.5	-0.65	0.35	0.0	-0.3
Etc...									

Table 6: Data Samples Format Table

Table 6 shows what the format of the extracted information would look like if it to be inserted into an excel data-sheet. From columns “Happy” to “Fear” the probability will consider all of the columns and be split respectively, as demonstrated in the second row of Table 6. The same applies to the columns from “Negative” to “Positive”, as sentiments are going to be measured separately from the emotions. As for the “compound” column, it is the sum of positive, negative & neutral scores which is then normalized between -1 (most extreme negative) and +1 (most extreme positive). The more the Compound score is closer to +1, the higher the positivity of the text, and vice versa; it can serve as a reference to the weight of the emotion behind the text.

positive sentiment: compound score >= 0.5
neutral sentiment: (compound score > -0.5) and (compound score < 0.5)
negative sentiment: compound score <= -0.5

Figure 15: Sentiments Using the Compound Score

As for the format that the merge function would follow, it would not differ too greatly from the format shown in Table 6, except that it will combine every excel file within the patient's folder in the order that they come in inside the folder. So if there were several excel files marking sessions inside a folder, then the merged file would organize the contents from the contents of the first file to the contents of the last file, as seen in Figure 16, and then become a new file, as shown in Figure 10.

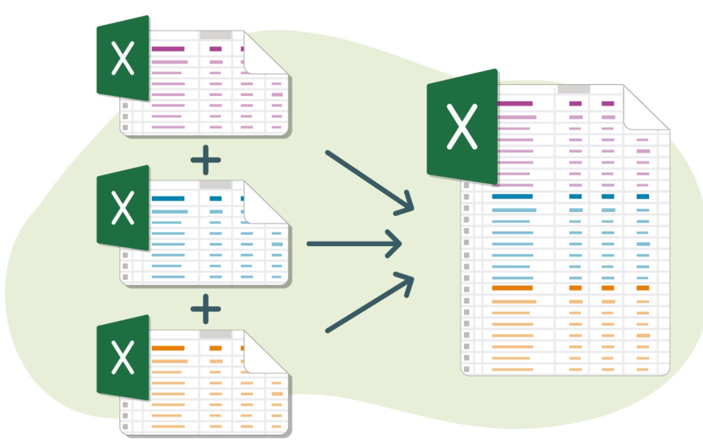


Figure 16: Merged Excel Files

3.3 Instruments and Procedures of Data Collection

The Pandas python library will be utilized to organize the analyzed data within a dictionary that has recorded the resulting text, sentiments, and emotions. The plan is to insert these values from their Pandas data frame format to an excel format, so that they may be used for data analysis.

Another instrument would be the Text2Emotion library that was mentioned earlier. This instrument will analyze the text that will be provided to it, and the end result will be converted into a JSON format and appended into an array. The information within that array will then be inserted into a selected excel file, as mentioned prior.

3.4 Data Analysis Procedures

As stated above, the information gathered from the user’s text will be recorded from a Pandas data frame into an excel format for data analysis. Through excel, the psychologist or patient will be able to use a familiar program to organize and graph the data at their leisure. The information would need to be graphed on a per row basis, so that all the emotions would be accounted for on the x-axis and measured numerically on the y-axis of a graph, without worrying about the user’s text as it would interfere with the generalized result.

For the analysis of the text, to calculate whether the patient suffers from symptoms of depression or anxiety, the analyzed emotions are what will allow for the clarification of the end result. So for the case of depression, the mean of the sadness, happiness, and compound score are all tested to come up with an appropriate evaluation befitting the conversation that was had with the chatbot.

$$f(s \vee h) = \frac{\sum_{i=0}^{n} (x_i)}{n}$$

Figure 17: Mean of Sadness/Happiness

What can be observed in Figure 17 is the mean used to measure both the sadness and happiness columns within the patient's session results where *n* is the size of the respective columns within these results. Both means will be used to measure the conditional probability for whether the user is depressed or not, with the compound score acting as the weight for the magnitude of the negative sentiment.

$$depressed = \begin{cases} f(s) & f(s) > 0.3 \\ f(h) & f(h) < 0.3 \\ f(s), f(h) & |(f(s) - f(h))| \geq 0.2 \end{cases}$$

Figure 18: Conditions for Depression Detection

Figure 18 shows the conditions that the program uses to determine whether the patient has depression or not. Using the previous figure as a reference, f(s) is the mean for the "Sadness" column while f(h) is for the "Happiness" column.

The same method can be applied for the detection of Anxiety using the program's analyzing feature. Although, it uses only the "Fear" column, along with the compound score, to measure whether signs of Anxiety exists within the patients texts.

$$anxiety = \begin{cases} f(a) & f(a) \geq 0.3 \\ f(c) & f(c) < 0.0 \end{cases}$$

Figure 19: Conditions for Anxiety Detection

For f(a) in Figure 19, it is the mean average of the Fear column. As for the f(c), it is the mean for the compound score column in which the weight is a negative value, so as to confirm that the overall sentiment of the conversation is within the context that it is intended for the measurement of an unfavorable condition. The results for both the Depression and Anxiety analysis will exist in the Analyse Page shown in Figure 7; the selected file will be graphed within the page using the ChartJS library and Jinja2, which is used alongside Flask.

CHAPTER 4

DATA ANALYSIS RESULTS

For the analysis of the results, patient Andy was selected. The conversation was as follows.

	Happy	Angry	Surprise	Sad	Fear	Negative	Neutral	Positive	Compound
Hello There	0	0	0	0	0	0	1	0	0
I usually like to sleep for long periods of time	0	0	0	0	1	0	0.76	0.24	0.36
Very often, mostly because I am unmotivated	0	0	0	1	0	0.32	0.68	0	-0.34
My lack of motivation stems from me feeling sad all the time	0	0	0	1	0	0.33	0.43	0.24	-0.36
Sleeping is one for now, though I feel sick looking at myself sometimes	0	0	0	1	0	0.23	0.77	0	-0.51

Table 7: Andy's Conversation Data

The chatbot would ask related questions based on the context of the conversation that it is currently having with Andy. The patient would tell the bot about his current happenings and the bot would respond in a manner befitting the circumstances.

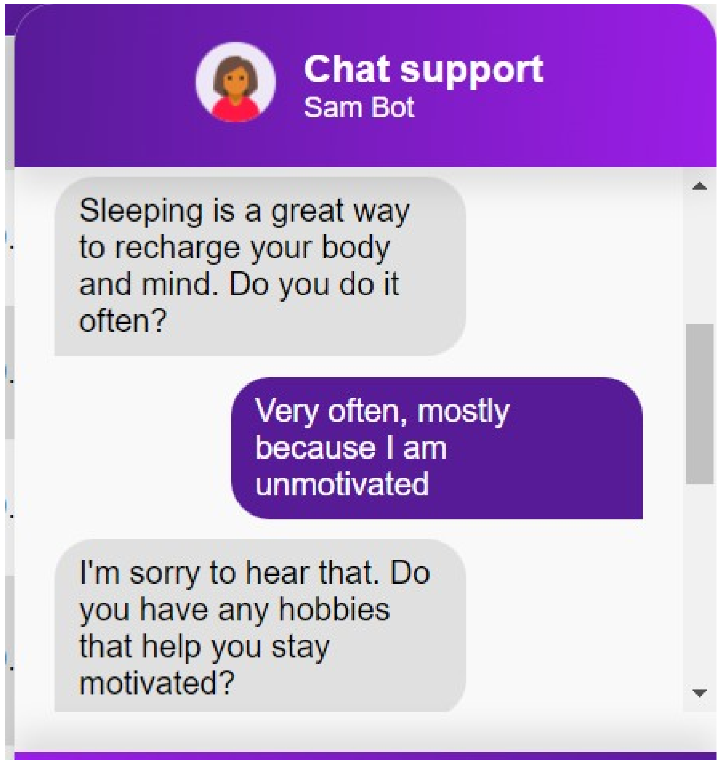


Figure 20: Chatbot Conversation Sample

As can be observed in Figure 20, the topic of conversation between the bot and the user in that moment was about hobbies. The user would tell the bot that their hobbies were to sleep, and the bot would replay appropriately. When the user would replay with a message that shows negative sentiments, the bot tries to ask questions that can allow the user to think from a more positive perspective while also sympathizing with the user.

Text	Happy	Angry	Surprise	Sad	Fear	Negative	Neutral	Positive	Compound
Hello there	0	0	0	0	0	0	1	0	0
I usually like to sleep for long periods of time	0	0	0	0	1	0	0.762	0.238	0.3612
Very often, mostly because I am unmotivated	0	0	0	1	0	0.324	0.676	0	-0.34
My lack of motivation stems from feeling sad all the time	0	0	0	1	0	0.331	0.429	0.239	-0.3612
Sleeping is one for now, though I feel sick looking at myself sometimes	0	0	0	1	0	0.231	0.769	0	-0.5106

Figure 21: Patient Results Preview Table

Before analyzing the whether the patient shows any signs of mental illness, the user can preview the results of the conversation they just had with the chatbot, as is shown in Figure 21. Once the results have been previewed, they can then be analyzed.

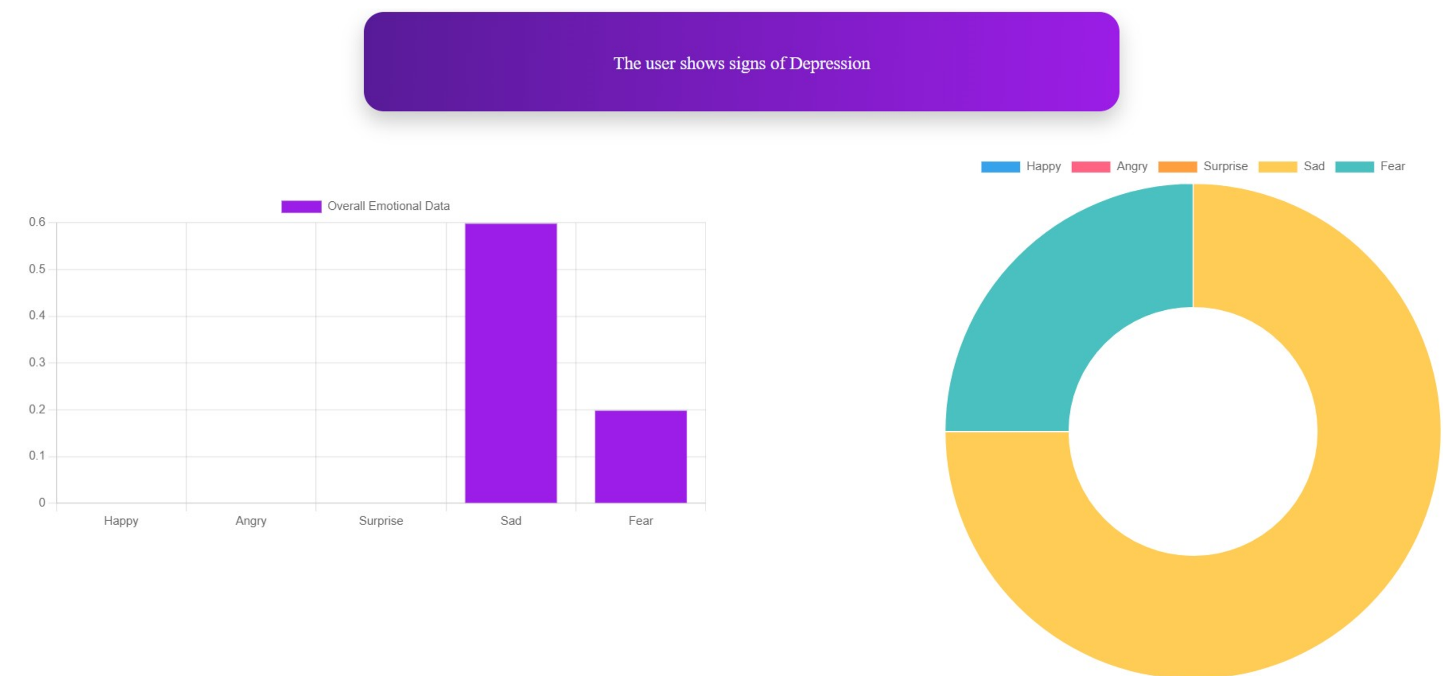


Figure 22: Analyzed Depression Results Page

Figure 22 shows that the analyzed results have concluded that the selected patient shows signs of depression. To arrive at that conclusion, the analyzer would perform the functions discussed in section 3.4, and since the results from the session's files meet the conditions for these functions, the analyzer was able to arrive at such a conclusion.

If Andy was to be tested for having symptoms of Anxiety, then the resulting conversation would be as follows.

	Happy	Angry	Surprise	Sad	Fear	Negative	Neutral	Positive	Compound
I have frightened myself out of many outings	0	0.5	0	0	0.5	0.33	0.67	0	-0.44
I can't live for the moment, like I'm having a midlife crisis	0	0	0	0	1	0.28	0.55	0.17	-0.38
Many attempts to discredit me	0	0	0	0	1	0	1	0	0
To have one's head on the chopping block	0	0	0	0	1	0.29	0.71	0	-0.44
To issue risk and have a horror of a system	0	0	0.5	0	0.5	0.49	0.51	0	-0.7
A system that does not let you take your life in your hands	0	0	0	0	1	0	1	0	0

Table 8: Andy's Anxiety Conversation Data

Just like with the previous result, the chatbot would ask relevant questions based on the patient's input and the result would be analyzed and later appended into an excel sheet. The fear column appears to hold a majority in the results, as this is a means to test the calculation discussed in section 3.4 regarding anxiety. The mean for the fear and the compound score exceed the threshold that has been placed to measure for anxiety, so when the file is analyzed, the result would be that Andy shows signs of anxiety.

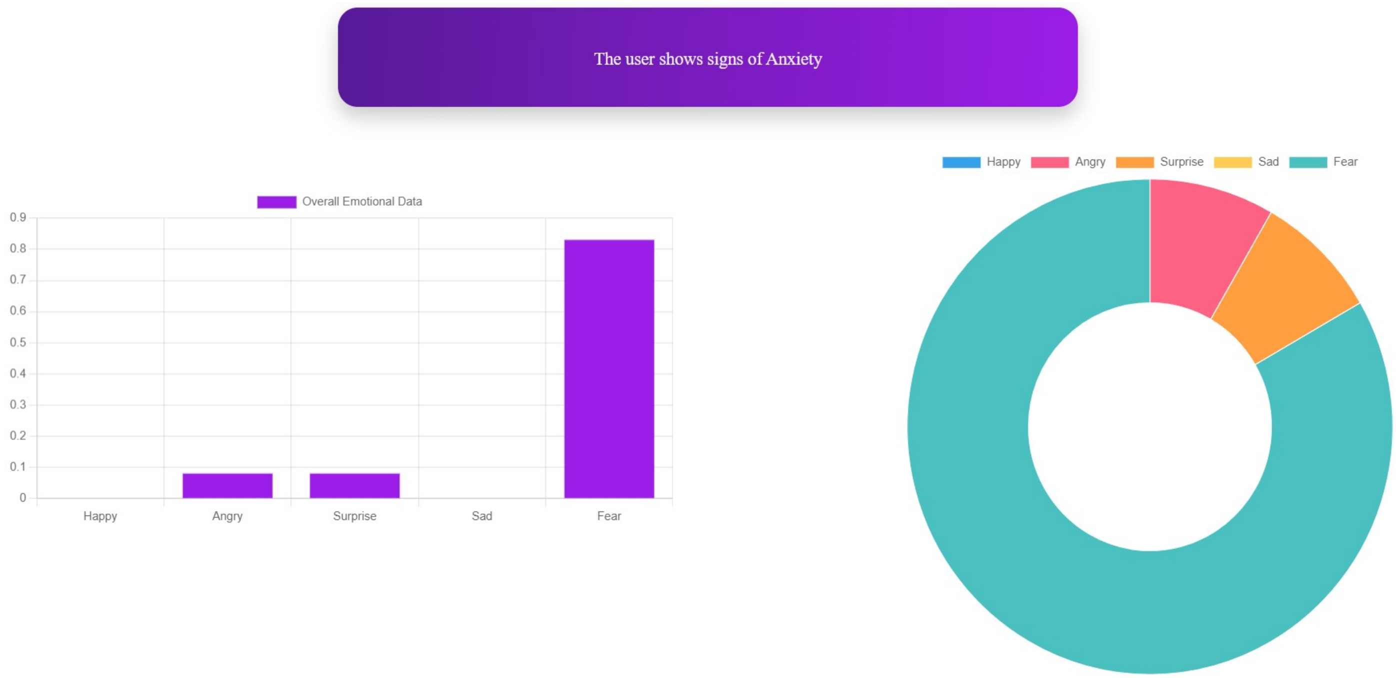


Figure 23: Anxiety End Result Andy

surmised, the end resulted with Andy being told they were showing signs of anxiety. On separate occasions of conversation with the bot, the bot would empathize with Andy.

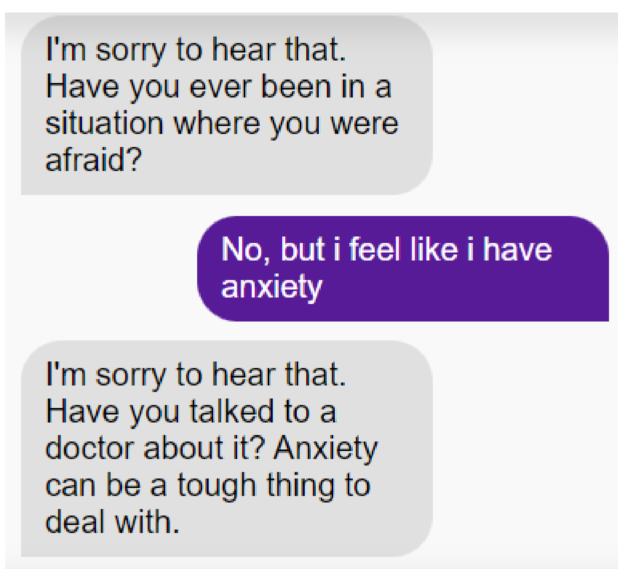


Figure 24: Bot Empathizing and Recommending Solutions

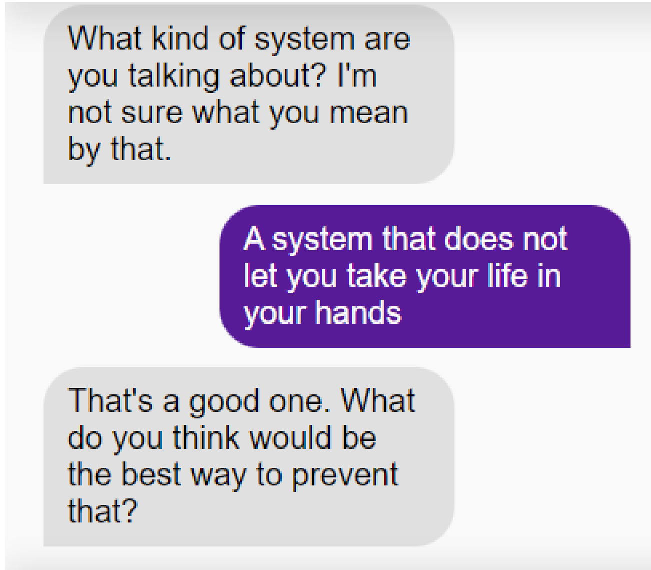


Figure 25: Bot Asking Insightful and Therapeutic Questions

Figure 24 and Figure 25 show a sample of what the is capable of in terms of pseudo-companionship through empathy and insight. Patient would describe their current troubles and the bot would either ask an insightful question or provide a recommendation, all showing the patient empathy for their situation.

Very often, mostly because I am unmotivated	0	0	0	1	0	0.324	0.676	0	-0.34
My lack of motivation stems from feeling sad all the time	0	0	0	1	0	0.331	0.429	0.239	-0.3612
Sleeping is one for now, though I feel sick looking at myself sometimes	0	0	0	1	0	0.231	0.769	0	-0.5106
It is absolute chaos	0	0	0	0	1	0.552	0.448	0	-0.5719
I am feeling tense from everything	0	0	1	0	0	0.348	0.435	0.217	-0.2263
I think I am having an anxiety attack	0	0	0	1	0	0.545	0.455	0	-0.5859
I live in fear of interaction	0	0	0	1	0	0.444	0.556	0	-0.4939
Only when chaos ensued in public	0	0	0	0	1	0.425	0.575	0	-0.5719

Figure 26: Andy's Merge Results Preview

The program is also capable of seeing the signs of Anxiety and Depression simultaneously. So if the user were to merge the folders, using a feature built into the program, then the program would analyze that the patient shows signs of both Depression and Anxiety. The file structure would look similar to Figure 10 once the feature is initiated, except that the merged file would not have a date and time naming convention, like with the other files within the patient's folder. Instead, all the data within the files in the patient's folder would be inserted into one file called "Merge {Patient Name}" inside the patient's folder. If the user were to choose to analyze said file then the preview would show what can be observed in Figure 26.

Information from several files can be observed in the preview, indicating a successful merge. Since the two sessions displayed in this section have shown different results, then the program would surmise that the user signs of both of these psychopathological conditions.

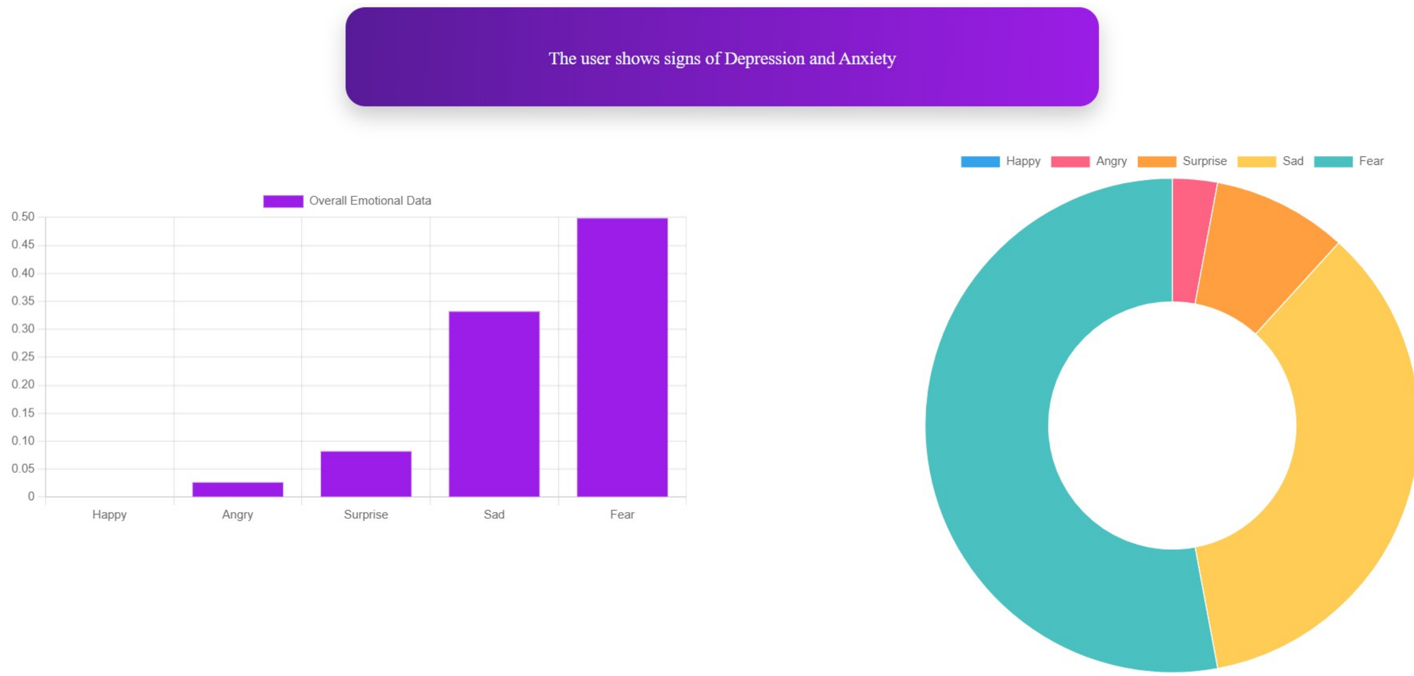


Figure 27: Merged Results

Figure 27 shows the result of the merged files, and as to be expected, the merged files have determined that the user has shown signs of both Depression and Anxiety. The more information is inserted into a file that is to be analyzed, the more detailed and accurate the results would be, since there would be a lot of data to work with.

CHAPTER 5

CONCLUSIONS AND IMPLICATIONS

5.1 Conclusions and Discussions

In conclusion, It is possible to predict what someone might be feeling through text. However, the accuracy varies based on the methodology used, since a deep learning approach, while costly in terms of hardware and intimidating in terms of complexity, is more accurate than a lexicon-based approach (Santosh et al., 2022). It was also discovered that the model for emotionally analyzing while conversing does have its benefits, since there are those that would confide in an AI according to Weizenbaum (1966), a Conversational AI offers a great gateway for escapism, since they are not judgmental (Asar, 2021). Additionally, if it were able to analyze the user's emotions accurately, then not only would the prevention of many negative effects from mental illness be decreased, but the creation of more life-like emotion feeling chatbots, through the implementation of these analyzers, would create a great many benefits in many sectors, be it psychology, gaming, or business.

5.2 Implications and Recommendations

The current implementation of an emotionally analyzing conversational AI is capable of engaging with the user, however it is still incomplete as its engagement does not yet feel personal. The pseudo-companionship will help bring out how people truly feel about something while attempting to evaluate their state of mind. With the assistance of a custom built Conversational AI, this implementation can be used to make AI better understand humans, meaning once it is perfected, it can be used to assist psychologists and psychiatrists or even make engaging game AI that can help people with little social experience gather said experience.

For the future, it would be recommended to use a learning-based approach for the AI, since it seems to be a far more effective method for creating human-like AI. This may create a blackbox of data, as stated prior, but that would not be an issue if the bot's behavior was monitored closely to make sure it does not deviate from the milestones that have been set.

REFERENCES

Acheampong, FA, Wenyu, C, Nunoo-Mensah, H. (2020).*Text-based emotion detection: Advances, challenges, and opportunities*. *Engineering Reports*; 2:e12189. https://doi.org/10.1002/eng2.12189

Aman, S., Szpakowicz, S. (2007). Identifying Expressions of Emotion in Text. In: Matoušek, V., Mautner, P. (eds) Text, Speech and Dialogue. TSD 2007. Lecture Notes in Computer Science(), vol 4629. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74628-7_27

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychological Association. (2021). *Apa Dictionary of Psychology*. American Psychological Association. https://dictionary.apa.org/psychopathology

Arjovsky, M. (2021, March 3). *Out of distribution generalization in machine learning*. arXiv.org. https://arxiv.org/abs/2103.02667

Asar, A. (2021, April 2). *Council post: Opening up to a robot? how mental health tech can help patients*. Forbes. https://www.forbes.com/sites/forbestechcouncil/2021/04/02/opening-up-to-a-robot-how-mental-health-tech-can-help-patients/

Bailey, M. M. (2019, August 31). *NRCLEX – The Library That Detects Emotions*. PyPI. https://pypi.org/project/NRCLex/

Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961) *An inventory for measuring depression*. *Archives of General Psychiatry, 4*, 561-571.

Biswas, S. (2021, September 29). *Advantages of deep learning, plus use cases and examples*. Width.ai. https://www.width.ai/post/advantages-of-deep-learning

Brezinka, V. (2012). *Computer games supporting cognitive behaviour therapy in children. Clinical Child Psychology and Psychiatry*, *19*(1), 100–110. https://doi.org/10.1177/1359104512468288

Brown, S. (2021, April 21). *Machine Learning, explained*. MIT Sloan. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Choi, S., Aizawa, K. Emotype: Expressing emotions by changing typeface in mobile messenger texting. *Multimed Tools Appl* 78, 14155–14172 (2019). https://doi.org/10.1007/s11042-018-6753-3

Coeckelbergh, M. (2016). *Chapter 9: The Future of Artificial Moral Agents*. In N. Bostrom & T. Ord (Eds.), *The Social Implications of Artificial Intelligence* (pp. 157). essay, Oxford University Press.

Dewey, C. (2020, September 18). *Replika and Virtual Friends.* https://www.conordewey.com/blog/replika-and-virtual-friends/

Ferro, R. (2019, December 31). *Psychopathology-Fer-Assistant: Because mental health matters*. GitHub. https://github.com/RodolfoFerro/psychopathology-fer-assistant

Bilquise G., Ibrahim S., Shaalan K. (2022). *Emotionally Intelligent Chatbots: A Systematic Literature Review*, *Human Behavior and Emerging Technologies.* https://doi.org/10.1155/2022/9601630

Gupta, A. (2020, September 13). *text2emotion Library*. GitHub. https://github.com/aman2656/text2emotion-library

Iwata, Y. (2021, October 24). *The boom in 'doing nothing'. Nikkei Asian Review.* https://asia.nikkei.com/Business/Consumers/The-boom-in-doing-nothing

Javaid, S. (2022, July 19). *Quick guide to datasets for Machine Learning in 2023*. AIMultiple. https://research.aimultiple.com/datasets-for-ml/

Juillion, P. (2019, October 4). *What is the difference between sentiment and emotion analysis?*. Studybuff.com. https://studybuff.com/what-is-the-difference-between-sentiment-and-emotion-analysis/

Justus, D., Brennan, J., Bonner, S., & McGough, A. S. (2018, November 28). *Predicting the computational cost of Deep Learning Models*. arXiv.org. https://doi.org/10.48550/arXiv.1811.11880

Li, D., Zhu, X., Li, Y., Wang, S., Li, D., Liao, J., & Zheng, J. (2021). *Enhancing emotion inference in conversations with commonsense knowledge*. *Knowledge-Based Systems*, *232*. https://doi.org/https://doi.org/10.1016/j.knosys.2021.107449.

Madhavi, K.R., Vooradi, S., Mounika, P., Yedlla, S., Tangudu, N. (2023). *Emotion Analysis Using Convolutional Neural Network*. In: Reddy, A.B., Nagini, S., Balas, V.E., Raju, K.S. (eds) Proceedings of Third International Conference on Advances in Computer Engineering and Communication Systems. Lecture Notes in Networks and Systems, vol 612. Springer, Singapore. https://doi.org/10.1007/978-981-19-9228-5_44

Majumder, S., Mondal, A. (2021). *Are chatbots really useful for human resource management?*. *International Journal of Speech Technology* 24, 969–977. https://doi.org/10.1007/s10772-021-09834-y

Mbunge, E., Jiyane, S., & Muchemwa, B. (2022). *Towards emotive sensory web in virtual health care: Trends, technologies, challenges and ethical issues*. *Sensors International*, *3*, 100134. https://doi.org/10.1016/j.sintl.2021.100134

Mehta, D., Siddiqui, M. F. H., & Javaid, A. Y. (2019). *Recognition of Emotion Intensities Using Machine Learning Algorithms: A Comparative Study*. *Sensors (Basel, Switzerland)*, *19*(8), 1897. https://doi.org/10.3390/s19081897

Mohammad, S. (2011, July 10). *NRC Word-Emotion Association Lexicon*. NRC Emotion Lexicon. https://www.saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm

Morante, R., & Blanco, E. (2021). *Recent advances in processing negation*. *Natural Language Engineering, 27*(2), 121-130. doi:10.1017/S1351324920000534

Nandwani, P., & Verma, R. (2021, August 28). *A review on sentiment analysis and emotion detection from text - social network analysis and Mining*. SpringerLink. https://link.springer.com/article/10.1007/s13278-021-00776-6

Park, S. C., & Kim, Y. K. (2020). *Anxiety Disorders in the DSM-5: Changes, Controversies, and Future Directions*. *Advances in experimental medicine and biology*, *1191*, 187–196. https://doi.org/10.1007/978-981-32-9705-0_12

Picard, RW. (1997). *Affect Detection: Uses, Contexts & Challenges*. https://affect.media.mit.edu/pdfs/95.picard.pdf

Purdy, M., Zealley, J., & Maseli, O. (2019, November 18). *The risks of using AI to interpret human emotions*. Harvard Business Review. https://hbr.org/2019/11/the-risks-of-using-ai-to-interpret-human-emotions

Rao, A. (2018, January 24). *Woebot- your AI cognitive behavioral therapist: An interview with Alison Darcy*. Medium. https://chatbotsmagazine.com/woebot-your-ai-cognitive-behavioral-therapist-an-interview-with-alison-darcy-b69ac238af45

Rudin, C., & Radin, J. (2019). *Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From an Explainable AI Competition*. Harvard Data Science Review, 1(2). https://doi.org/10.1162/99608f92.5a8a3a3d

S. Kaltwang & F. Maurer, (2006). *Automatic Analysis of Facial Expressions and Body Language: State of the Art*, in *Advances in Visual Computing VII*, vol. 331 of Lecture Notes in Computer Science (Ed.), S. Marcel, M. Santana, E. Vicente, and N. García.

Santosh Kumar Bharti, S Varadhaganapathy, Rajeev Kumar Gupta, Prashant Kumar Shukla, Mohamed Bouye, Simon Karanja Hingaa, Amena Mahmoud, "Text-Based Emotion Recognition Using Deep Learning Approach", *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2645381, 8 pages, 2022. https://doi.org/10.1155/2022/2645381

Schimelpfening, N. (2022, April 27). *What you need to know about the DSM-5 updates on depression*. Verywell Mind. https://www.verywellmind.com/dsm-5-and-diagnosis-of-depression-1066916

Stomp, M. (2018). *Animals' Ability To Recognise Themselves in Mirror Tests Is Affected By Whether They Live With Humans* (Vol. 8). Scientific Reports.

Tempelaar D, Rienties B, Nguyen Q (2020). *Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application*. PLoS ONE 15(6): e0233977. https://doi.org/10.1371/journal.pone.0233977

Weizenbaum, J. (1966). *ELIZA—A computer program for the study of natural language communication between man and machine. Communications of the ACM*, 9(1), 36–45. https://doi. org/10.1145/357980.357991

Weston, J., & Shuster, K. (2021, July 16). *Blender Bot 2.0: An open source chatbot that builds long-term memory and searches the internet*. Meta AI. https://ai.facebook.com/blog/blender-bot-2-an-open-source-chatbot-that-builds-long-term-memory-and-searches-the-internet/

Witek, M. (2014). Linguistic underdeterminacy: A view from speech act theory. *Journal of Pragmatics*, *76*, 15–29. https://doi.org/https://doi.org/10.1016/j.pragma.2014.11.003.

Wolff, J., & Stroulia, E. (2005). Conversational Agents Facilitating Verbal & Nonverbal Communication. In J. R. Brown (Ed.), *Embodied Conversational Agents* (pp. 30–48). essay, Springer Netherlands.

Yam, C. (2020, March 15). *Emotion detection and recognition from text using Deep Learning*. CSE Developer Blog. https://devblogs.microsoft.com/cse/2015/11/29/emotion-detection-and-recognition-from-text-using-deep-learning/

APPENDICES

APPENDIX A
User Documentation for Setting Up and Running The Program

In this section is the setup for the program mentioned in this paper. The method of setting up can also be found in the README.md file in the project directory.

The first step is to run the virtual environment.

- Open the command prompt or terminal

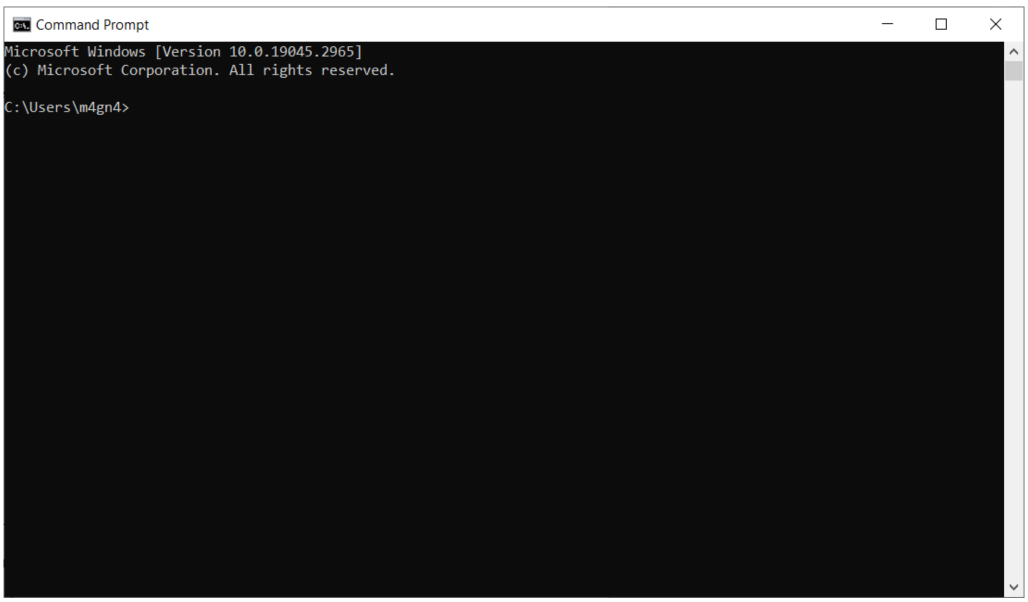


Figure 28: Default State Command Prompt

- Confirm that the Python programming language is installed and in the PATH directory by typing "python –version"

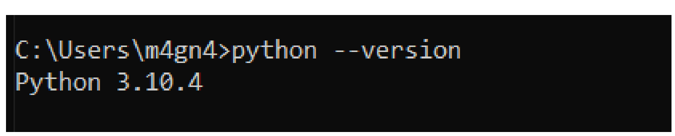


Figure 29: Python Version Confirmation

- If python is installed, navigate to the project directory via "cd {Project Directory}"

```
C:\Users\m4gn4>cd C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2

C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>
```

Figure 30: Project Directory Navigation

- Type "venv/Scripts/activate" to run the virtual environment in the Visual Studio Code terminal, and ".\venv\Scripts\activate" to run it in the Command Prompt

```
C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>.\venv\Scripts\activate

(venv) C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>
```

Figure 31: Command Prompt Virtual Environment

```
PS C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2> venv/Scripts/activate
(venv) PS C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>
```

Figure 32: Visual Studio Code Terminal Virtual Environment

The (venv) that is seen in both Figures 31 and 32 is an indication that the Command Prompt/Terminal is currently running the virtual environment. The next step is to install the requirements within the virtual environment to properly run the program. So perform the following steps with the command prompt or terminal still open.

- In the command prompt/terminal, type “pip install -r requirements.txt” to install the required libraries

```
(venv) C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>pip install -r requirements.txt
```

Figure 33: Requirements Installation with pip

These requirements will be installed within the virtual environment, meaning that if one of the libraries within the requirements were to be used outside the virtual environment, then the program would give an error.

- Once the installation is complete, type python in the command prompt/terminal. This is for the installation of “punkt” for the “nltk” library

```
(venv) C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>python
Python 3.10.4 (tags/v3.10.4:9d38120, Mar 23 2022, 23:13:41) [MSC v.1929 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

Figure 34: Running Python Scripts Command

- The command prompt will display “>> ” indicating that the python script is being run
- In that “>> ” type “import nltk” then press enter, so that the script would register the import, then type “nltk.download(‘punkt’)”

```
>>> import nltk
>>> nltk.download('punkt')
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\m4gn4\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
True
```

Figure 35: NLTK punkt Package Download

The step that comes after the installation is running the program. Therefore, while still in the command prompt/terminal perform the following.

- Within the ">> " type "quit()" to leave the python script and return to the project directory of the virtual environment

```
True
>>> quit()

(venv) C:\Users\m4gn4\Desktop\FIU Courses Files\Thesis\Final Source Code 2>
```

Figure 36: Returning To Virtual Environment From Python Script

- Run the program by typing "python app.py", this will open a new flask server

If all the steps have been followed correctly thus far, then the Command Prompt/Terminal should display the following information shown in Figure 33.

- The flask server's port will be shown in the command prompt/terminal so insert that port into a web browser of choice (Chromium based would be recommended), or do "ctrl+click" on the displayed port's URL (feasible if these steps were being followed on the Visual Studio Code terminal)

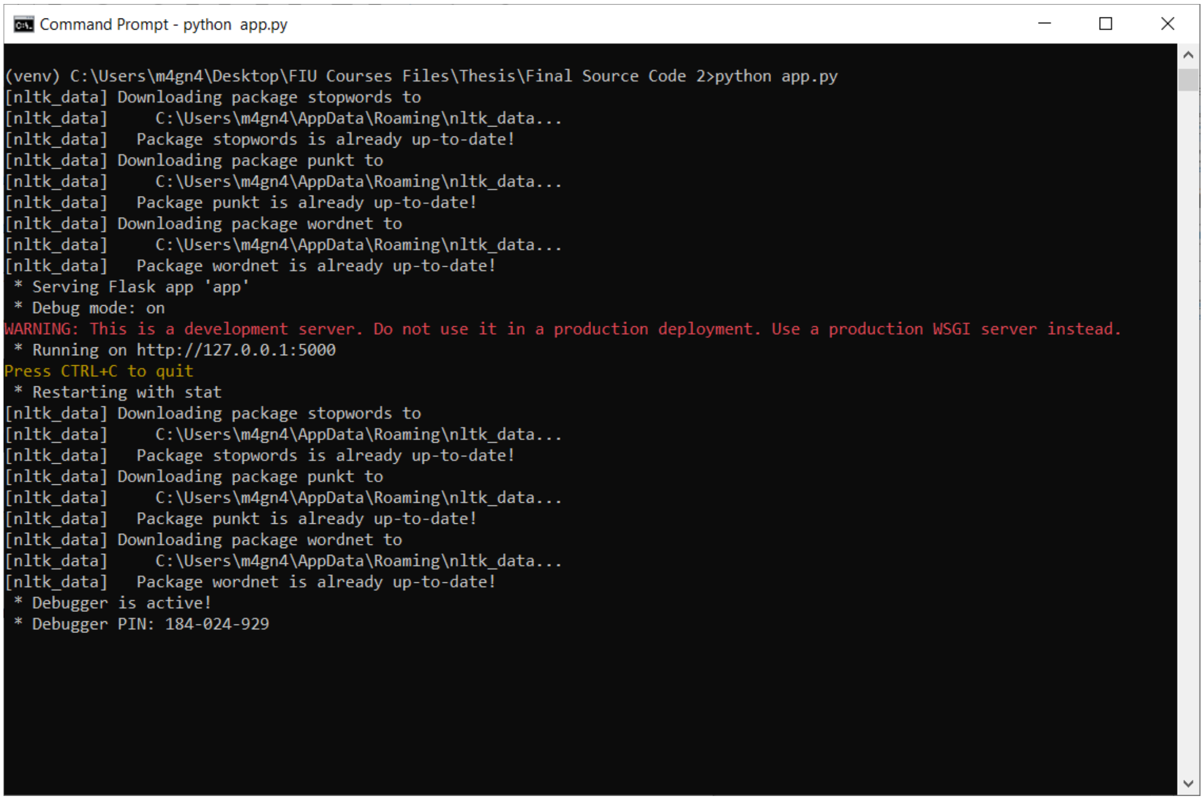


Figure 37: Command Prompt for Running the Flask App

APPENDIX B
User Guide for Using The Program

If the setup in APPENDIX A was done correctly, the user should be met with a landing page that asks the user to select a patient or add a new one.

Figure 38: Landing Page on Port 5000

As can be observed in Figure 38, the landing page is on a local server on port 5000. The user is met with a dropdown selection and a button that asks them to add a new user. If the dropdown is selected

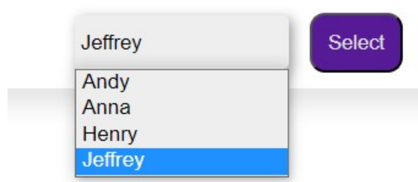


Figure 39: Dropdown Patient Selection

Then options will appear for existing patients. These patients are derived from the emotion_data folder in the project directory

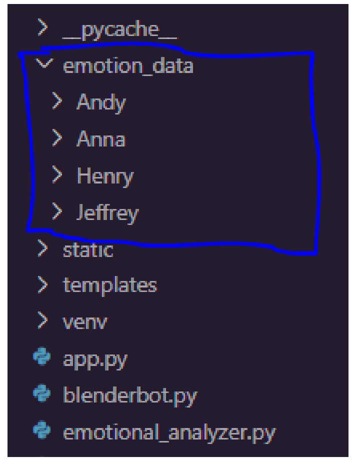


Figure 40: Emotion Data Patient Folders

If the user were to press the "Add New" button, they will be met with a modal that asks for the new patient's name

Figure 41: Add New Patient Modal

For this example, the name Jeremy will be used as a placeholder, and if the user were to press the "Add" button

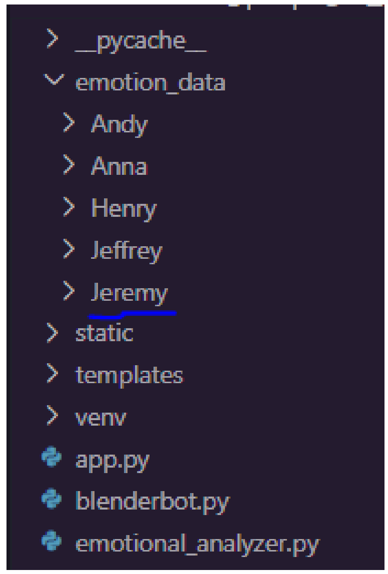


Figure 42: New Patient Jeremy in Emotion Data Folder

then patient Jeremy will be added to the emotion_data directory as a new record, but since the user did not start a session with Jeremy yet, Jeremy's directory is empty.

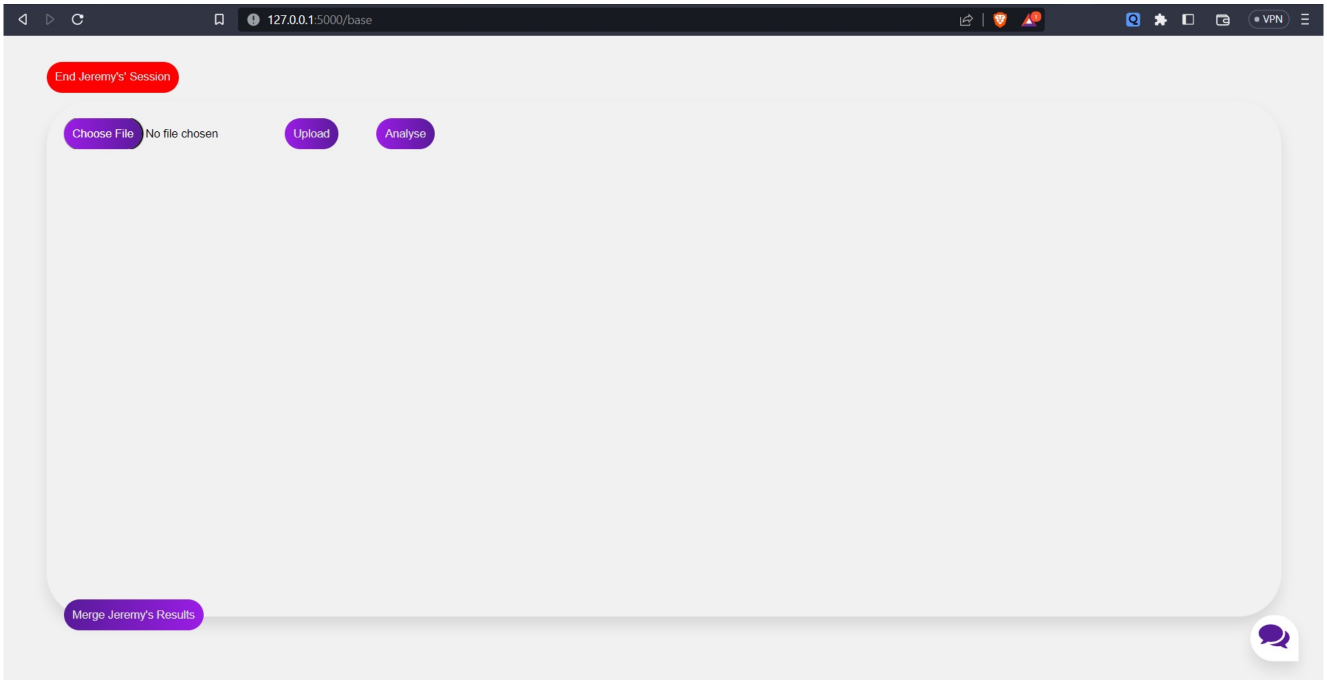


Figure 43: Jeremy in Base Page

Additionally, pressing the "Add" button to add a new patient will navigate the user to the Base URL, in accordance to the structure shown in Figure 7. If the user were to press the "End Jeremy's Session" button, then they will be navigated back to the landing page, and should the user select a patient from the dropdown

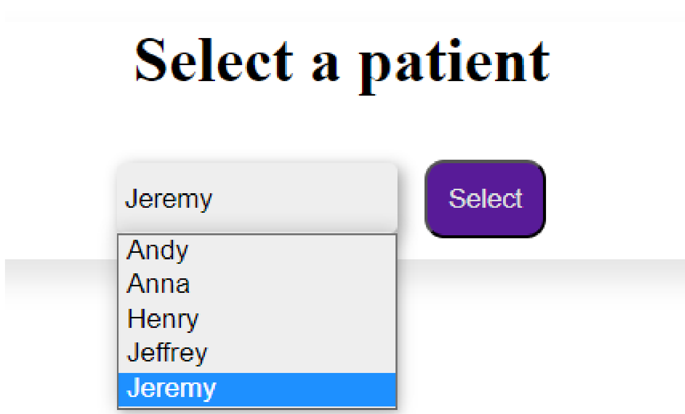


Figure 44: Jeremy as a New Dropdown Selection

then Jeremy can be seen as part of the users that are selectable.

Starting a session with Jeremy is simple, as it only requires selecting Jeremy, pressing the Chat icon on the bottom right corner, and starting a new conversation with the chatbot

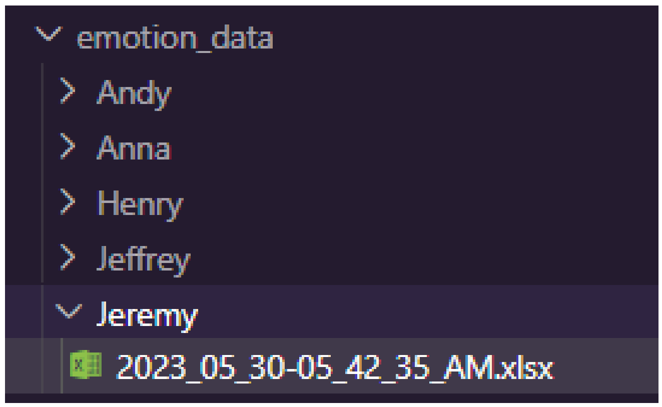


Figure 45: Jeremy New Session Start

By initiating a conversation with the chatbot, a new session has been inserted into Jeremy's folder, all while the excel file is named the exact time and date that the session started. The session remains active until the user either refreshes, to start a new session with the same patient, or leaves, be it to end the session or analyze it.

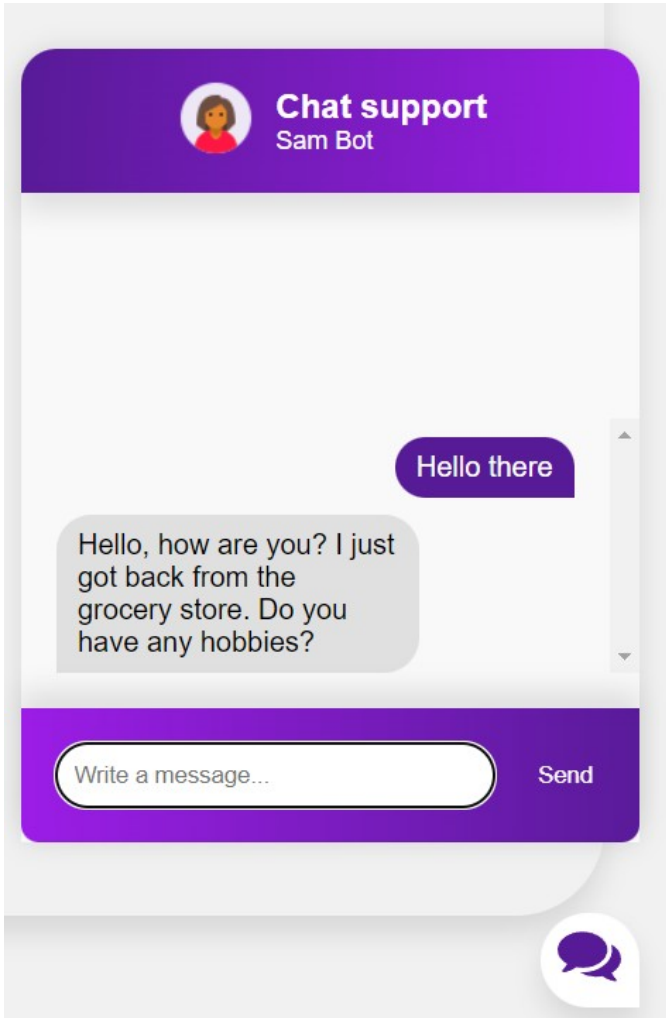


Figure 46: Starting New Session By Engaging With Bot

After having the conversation with the chatbot, the user can select the file for the current session using the Choose File button

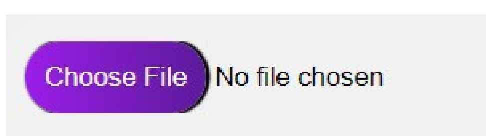


Figure 47: Choose File From Folders Button

Upon pressing the button, a popup prompting the user to select which file they would like to preview will appear.

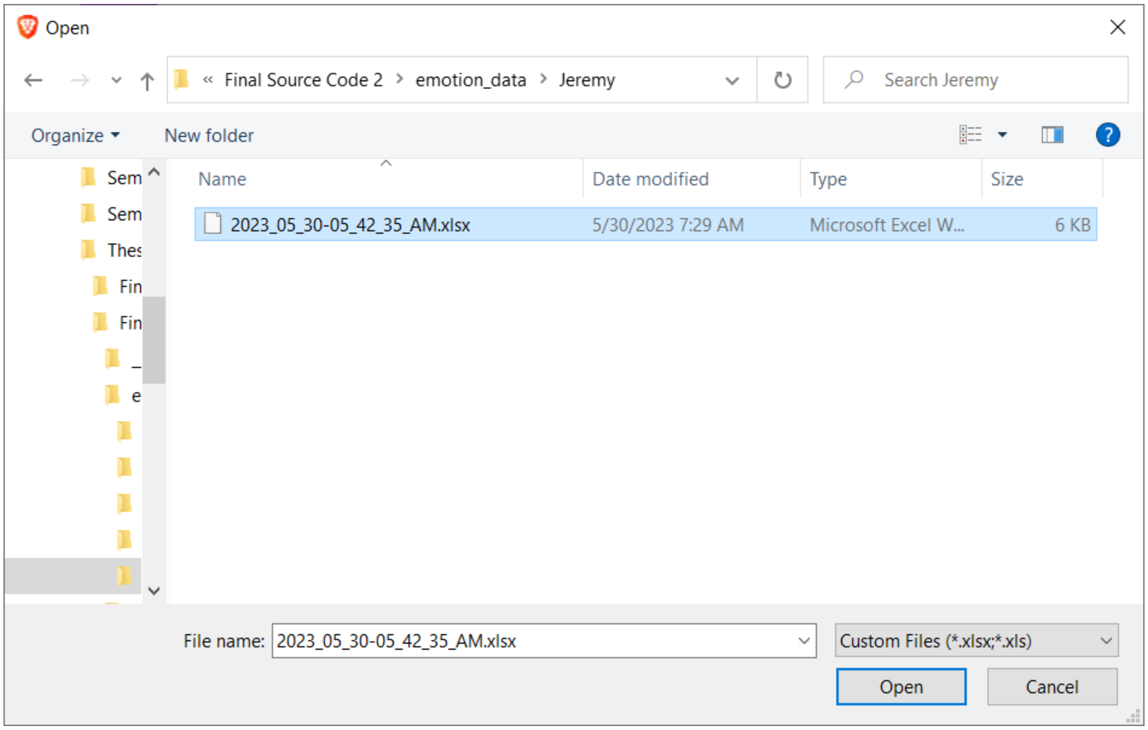


Figure 48: Selecting Recent Session File

The user will select the current session and press open. Once the file has been selected, the next step would be to press the upload button to confirm the selection and begin previewing the conversation as well the emotions and sentiments that have been analyzed.

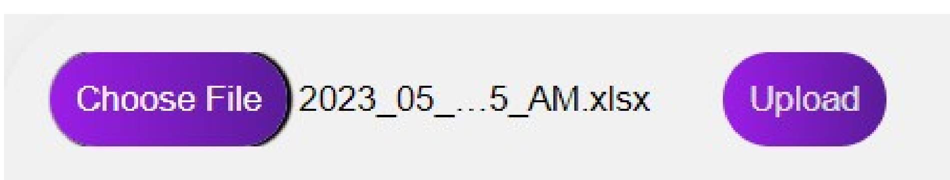


Figure 49: Selected Session File and Upload Button

Text	Happy	Angry	Surprise	Sad	Fear	Negative	Neutral	Positive	Compound
Hello there	0	0	0	0	0	0	1	0	0
My hobbies are eating outside with my friends and playing video games	0.33	0	0	0	0.67	0	0.671	0.329	0.5994
Mostly story driven games, it is like being part of a book except you can interact with the book	0	0	0	0	1	0	0.872	0.128	0.3612
Or rather, you are the fictional character	0	0	0	0	1	0	1	0	0
I was talking about the book	0	0	0	0	0	0	1	0	0
A book about dungeons and dragons	0	0	0	0	0	0	1	0	0

Figure 50: Preview Jeremy's Conversation and Analyzed Emotions

This preview was demonstrated in Section 4, and it shows the results of the analysis for the emotional data given the user's text. The preview can be updated with new information if the user were to continue its conversation with chatbot and re-upload the file by going through the steps mentioned prior, about selecting the file for the current session and uploading it.

Now that Jeremy is done with their conversation, all that is left is to press the "Analyse" button, so that they may be taken to the "Analyse" page where the end result is calculated and graphed.

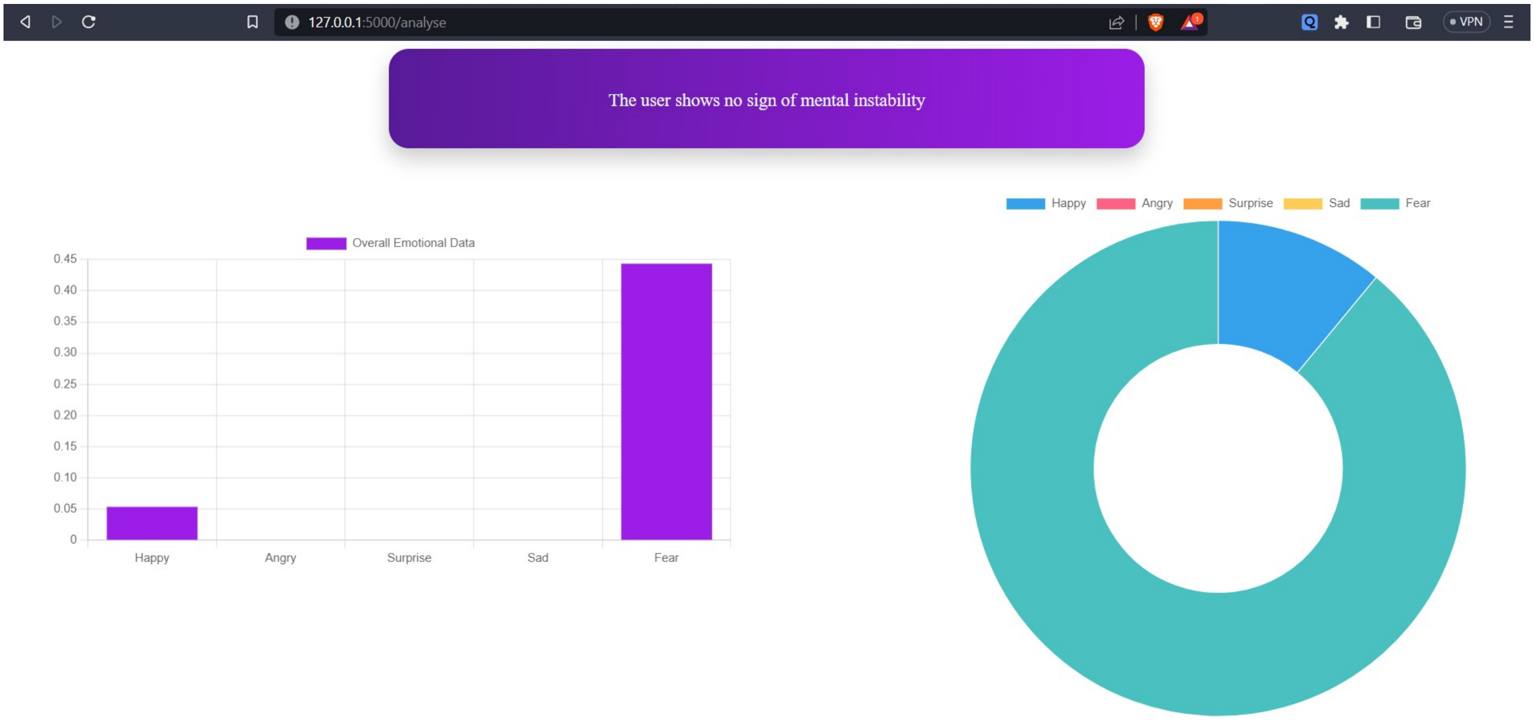
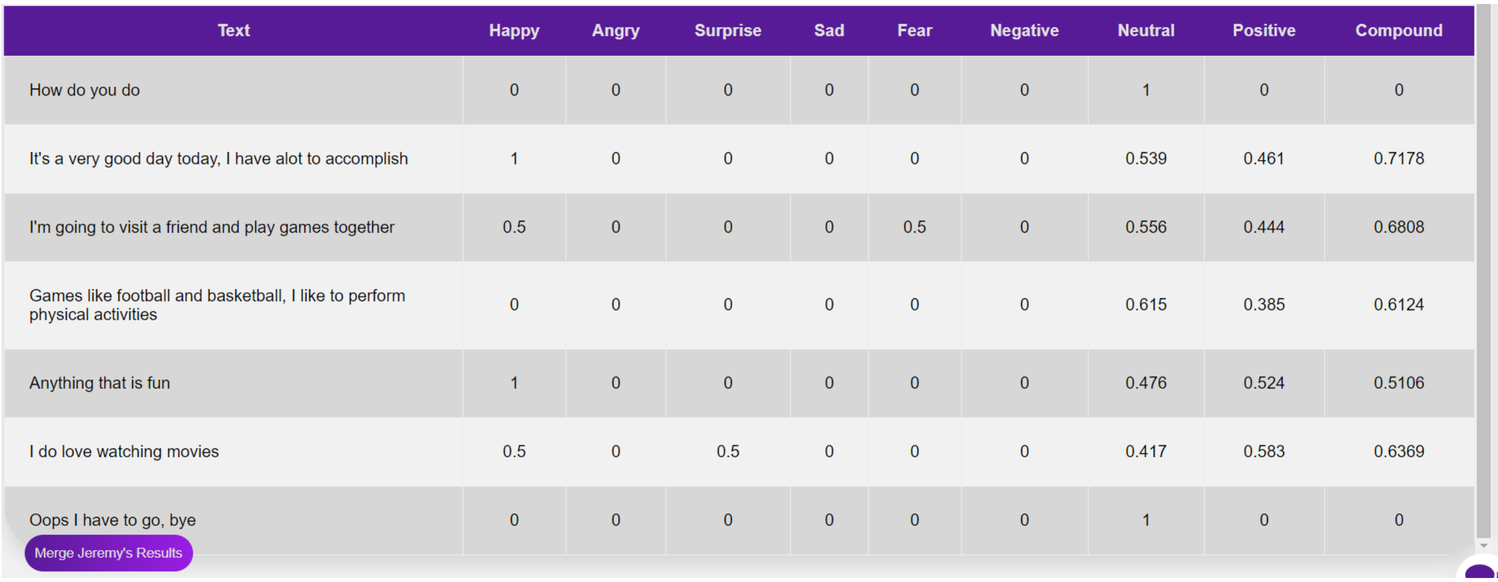


Figure 51: Jeremy Result of First Session

As can be observed in Figure 51, Jeremy shows no signs of mental illness; he is neither suffering from symptoms of depression nor symptoms of anxiety. The end result of the graphs also shows which emotions were more prominent than others based on the size of the information found in the session's excel sheet.

Since the session's data size was not that big, it is possible to expand the size of data by merging previous or post sessions. This program contains a feature that allows for the merging of all the sessions of a selected patient, which can be utilized to gather an overall insight into the patient's mental condition.



Text	Happy	Angry	Surprise	Sad	Fear	Negative	Neutral	Positive	Compound
How do you do	0	0	0	0	0	0	1	0	0
It's a very good day today, I have alot to accomplish	1	0	0	0	0	0	0.539	0.461	0.7178
I'm going to visit a friend and play games together	0.5	0	0	0	0.5	0	0.556	0.444	0.6808
Games like football and basketball, I like to perform physical activities	0	0	0	0	0	0	0.615	0.385	0.6124
Anything that is fun	1	0	0	0	0	0	0.476	0.524	0.5106
I do love watching movies	0.5	0	0.5	0	0	0	0.417	0.583	0.6369
Oops I have to go, bye	0	0	0	0	0	0	1	0	0

Figure 52: Preview of Jeremy's Second Session

With the preview of Jeremy's second conversation, the user is able to press the "Merge Jeremy's Results" button on the bottom left of the figure in order to merge all the files in Jeremy's directory, including the most recent one.

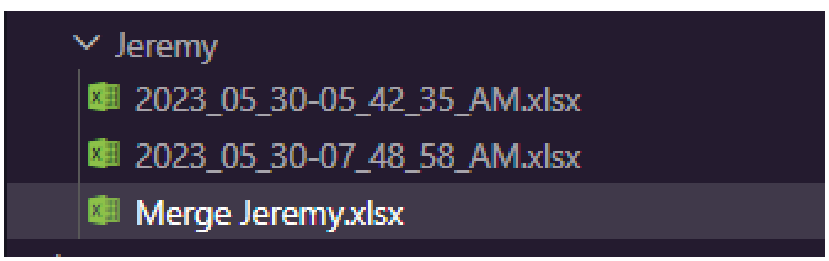


Figure 53: Generated Merge File for Jeremy

The files merge into one and create a much larger data-sheet to be analyzed.

Or rather, you are the fictional character	0	0	0	0	1	0	1	0	0
I was talking about the book	0	0	0	0	0	0	1	0	0
A book about dungeons and dragons	0	0	0	0	0	0	1	0	0
How do you do	0	0	0	0	0	0	1	0	0
It's a very good day today, I have alot to accomplish	1	0	0	0	0	0	0.539	0.461	0.7178
I'm going to visit a friend and play games together	0.5	0	0	0	0.5	0	0.556	0.444	0.6808
Games like football and basketball, I like to perform physical activities	0	0	0	0	0	0	0.615	0.385	0.6124
Anything that is fun	1	0	0	0	0	0	0.476	0.524	0.5106

Figure 54: Preview of Jeremy's Merged Files

It can be established from this preview that the files have been successfully merged, since the last text said in the first session can be seen right above the first text from the current session. Upon pressing the “Analyse” button

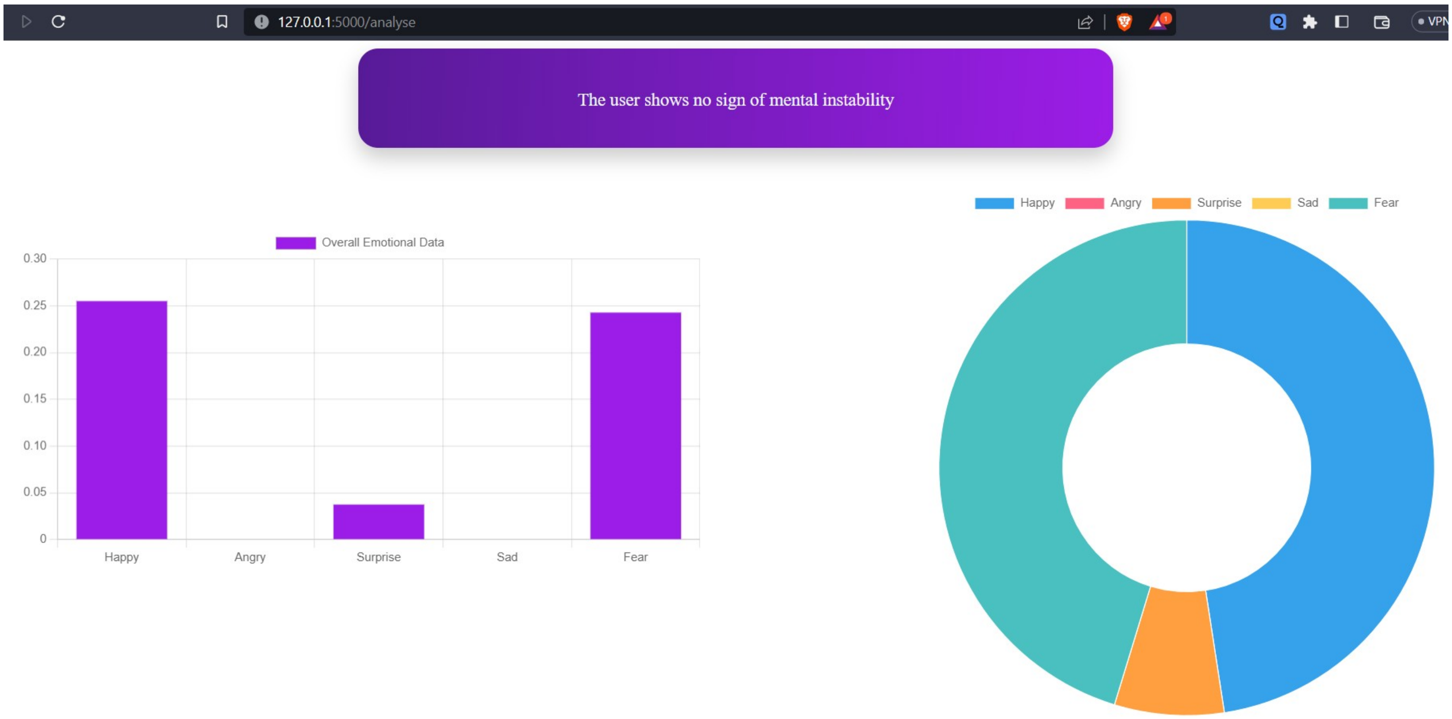


Figure 55: Jeremy's Merged Data Results

the results display a lot more information than with Jeremy's first session.