

SAMRAT  
SAHOO

COMPUTER SCIENCE

# ORIGINAL WORK PROPOSAL

INDEPENDENT STUDY  
MENTORSHIP

---

DATE

OCTOBER 1, 2020

PROPOSED BY

SAMRAT SAHOO

# Original Work Proposal

October 1, 2020

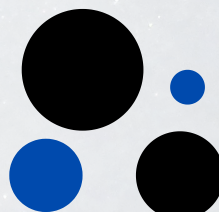
## Problem & Solution Statement

Everyday, more than 264 million people globally suffer from depression. Additionally, another 40 million suffer from anxiety. Tackling such a massive problem on an individual basis can be one of the most daunting challenges in mental health humanity has faced in the past few decades. It can cost billions and the feasibility of tending to the needs of each and every person is near impossible.

The product I am proposing offers a scalable, efficient, and cost-saving approach that leverages one of the most prominent social media platforms today: Twitter. For the original work, I plan to create a Twitter bot that utilizes the power of linguistics and artificial intelligence to both identify mental health issues such as depression and anxiety while also performing a therapy through talking with the user.

## Project Phases

This project has four steps to from start to completion. The first step would be data collection. With any machine learning project, it is necessary to collect hundreds of thousands of pieces of data. This project will web scrape tweets using common words that indicate a mental illness such as depression as well as regular tweets with no indication of depression. The second step is creating the depression classification model. This phase will use the data from the first step to help the computer learn what depressing text looks like. Using a neural network, the computer will be able to classify tweets into depressing and regular and then this model will then be saved. The third phase of this project includes the chat bot feature which will direct message a user to help the user with his or her problems. Using a recurrent neural network, the bot can generate appropriate messages based on the user's responses. The final phase of this project connects all prior aspects through the Twitter API in order to create an actual Twitter account that functions like a bot.



SAMRAT  
SAHOO

## Methodology

Natural language processing is still a relatively new field for me and I do not have much background knowledge in linguistics. To overcome this barrier, the first necessary step would be research. Here, it will be important to gain an understanding over natural language processing specific technologies while understanding how machine learning technologies are used in natural language processing. It may also prove useful to do some research in plain English to understand the language better. This research will be helpful in the sense that I will have an intuitive understanding of the code I write.

The second necessary step would be to explore the technical aspects of different computer science libraries so I can start writing actual code. This includes different natural language processing libraries such as natural language toolkit (NLTK), a machine learning library like PyTorch, or the Python library for Twitter, Tweepy. By dedicating some time to learn these libraries, I can more skillfully use them in order to increase my breadth of functionality of my project and produce a more refined product.

After gaining a better understanding of natural language processing concepts and learning the different aspects of these libraries, the third necessary step would be the development process. The development process would complete the four phases as outlined in the project phases section. To document this work, I plan on streaming my programming sessions on YouTube or creating detailed documentation over the product at the end of the development process.

## Extensibility

This project has limitless opportunities for extensibility. One avenue of extensibility is making it multi-platform. Right now the project is planned to solely run on Twitter; however, it has the potential to run on a wide array of social media platforms including Instagram and Reddit. Additionally, this project relies solely on natural language processing to compute all tasks; however, many posts include pictures. Using computer vision, the bot would also be able to analyze posts with pictures. Because of such extensibility, this project serves as a foundation to solving the world's mental health crisis

