

ISSR-3 Evaluation Test

**AI-Powered Behavioral Analysis for Suicide Prevention,
Substance Use, and Mental Health Crisis Detection with
Longitudinal Geospatial Crisis Trend Analysis**

20.03.2024

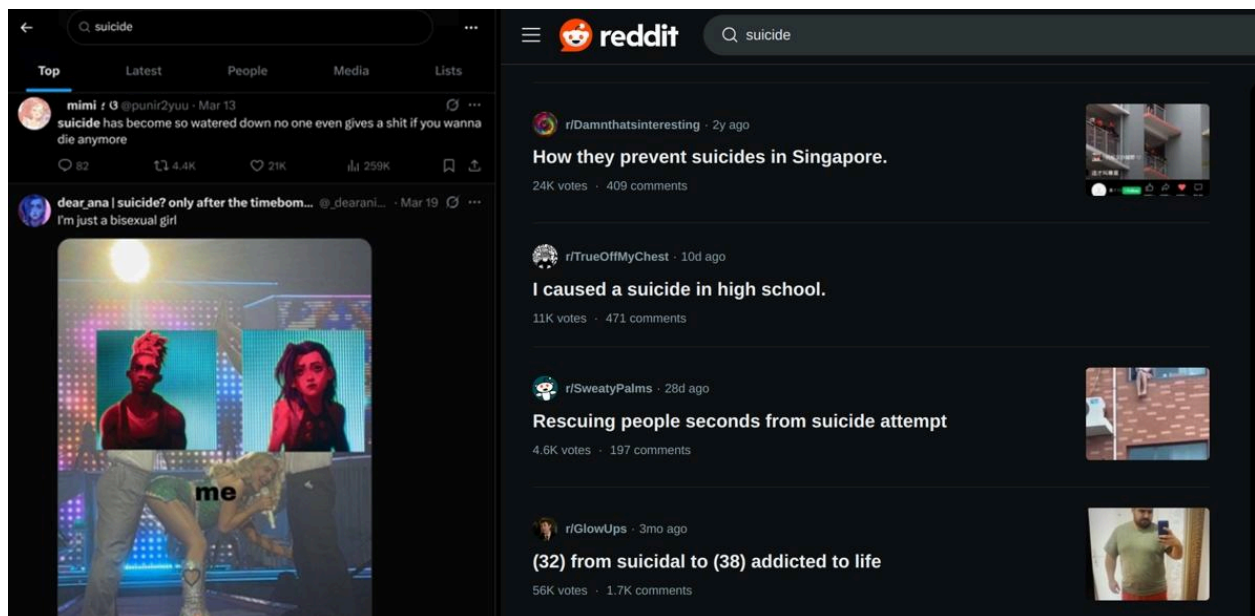
Haleel Sada N P

Abstract

This project focuses on extracting, analyzing, and visualizing mental health crisis-related discussions from social media platform Reddit. Posts related to mental health distress, substance use, and suicidality were collected based on predefined keywords using APIs. The data was preprocessed and sentiment classification was performed. Additionally, I have tried to identify regional crisis patterns, with results visualized. I hope these insights will enable me to contribute toward making the world a better and more supportive place for those in crisis...

Introduction

The objective of this project was to extract, analyze, and visualize crisis-related discussions from social media platforms like Twitter (X) and Reddit. However, instead of simply searching for predefined keywords, which often return news articles, past events, or even unrelated content like memes, I took a different approach. My goal was to make this project a leading indicator rather than a lagging one. Rather than analyzing past incidents, I focused on identifying individuals who might currently be in distress, allowing for potential early intervention.






Fig 1. Example of the results of search query, X on left side and reddit on right. 1 random video and one meme and 1 irrelevant post.

For this reason, I decided to work exclusively with Reddit, specifically subreddits like r/SuicideWatch and r/depression. These communities contain deeply personal and honest discussions, something that cannot be easily captured through surveys or news observation. Reddit's anonymity provides users with a safe space to express their emotions, making it a valuable source of unfiltered mental health data. I scraped recent posts from these subreddits, cleaned and processed the data, and stored it for further analysis.


For sentiment classification, I used **VADER**, as it performed better than TextBlob in classification. For risk categorization, I fine-tuned a **BERT-base-uncased** model by adding an additional classification layer and training it on a high-quality dataset sourced from Kaggle. This allowed for better identification of posts that indicated different levels of distress.

```
I just want to commit suicide : Negative
I am so happy today : Positive
I am sad but I am ok : Nuetral
Enter your post: I am feeling really bad now, maybe I will kill myself.
sentiment: Negative
```

Fig 2. Vader input and sentiment output

The final task—geolocation tracking—was a bit challenging. Unlike crisis situations involving physical injuries or accidents, where location-based help is more feasible, suicidal individuals rarely post their location directly. A phrase like "Need help in Austin" is uncommon because people in such distress often do not expect external physical assistance. To address this, I took a different route: analyzing the post history of users to determine if they were active in local city-based subreddits. While this approach occasionally provided clues about their location, many users create anonymous accounts solely to discuss their struggles, making identification difficult. Without access to IP addresses (which Reddit does not provide), precise geolocation remains a major challenge.

Despite these limitations, I was able to generate a meaningful visualization of distress patterns based on the available data. With this small effort, I hope to contribute something meaningful to help those who struggle every day through life's hardships.



Additionally I tried to take a look at what subreddits are these people in crisis mostly visits. They were mentioned in the Jupyter notebook. From that list if I remove the subreddits the population visits and the one I have used for this study, few ones are provided here. In ascending order of count.

genz, ocd, lonely, teenagers, soberlifeprotips, redditorsinrecovery, opiatesrecovery, stopdrinking

It is notable that most people who suffer from depression or drug abuse seek help through this subs r/stopdrinking, r/opiatesrecovery, r/redditorsinrecovery. While the subreddits I have used (r/depression, r/selfharm) and all are safe places where people can vent, these subreddits listed with the most counts are for seeking help and support, and it is notable that r/teenagers came before r/adulthood which is concerning.

It is amusing that the location maps show patterns, most people in crisis are from US, India or Europe. While this location map is nowhere near being accurate it is still amusing how more cases are accumulated in US states. It could also be biased considering the fact that reddit is not being used worldwide and people in developed countries use it more unlike platforms like Instagram or Facebook.

Images in the next page

More Results

	ID	Timestamp	Title	Content	Content_row	Author	Likes	Comments	Subreddit
0	doqwow	1.572361e+09	Our most-broken and least-understood rules is ...	understand people reply immediately op invitat...	We understand that most people who reply immed...	SQLwitch	2363	177	depression
1	1frqlk0	1.727565e+09	Regular check-in post, with information about ...	welcome rdepressions checkin post place take m...	Welcome to /r/depression's check-in post - a p...	SQLwitch	42	263	depression
2	1jie7of	1.742776e+09	I have cried so much, im tired	name josh im im lot pain friends spend time th...	My name is Josh, im 33 and im in a lot of pain...	CucumberCultural3760	126	32	depression
3	1ji2co9	1.742745e+09	I've spent 90% of my freetime in bed since I w...	title says ive bedrotting since spend time lay...	As the title says I've been bedrotting since I...	hylskrik	292	27	depression
4	1jiawlg	1.742767e+09	Being an underachiever with depression is a do...	least overachievers peoples respect theyre see...	At least overachievers have other people's res...	Nitrogen70	88	7	depression
5	1jik7b0	1.742797e+09	Every Time I Fall Asleep I'm Hoping my eyes Ne...	im life virtual im unemployed keep getting rej...	I'm 18 about to 19 and once I am My life is Vi...	ValTorni	19	2	depression
6	1jicxob	1.742772e+09	IT REALLY DID GET BETTER	year bet rotting way scared talk woman really ...	Before this year I was bet rotting being way t...	plaguepsycho419	55	29	depression

Fig 3. Collected reddit post data

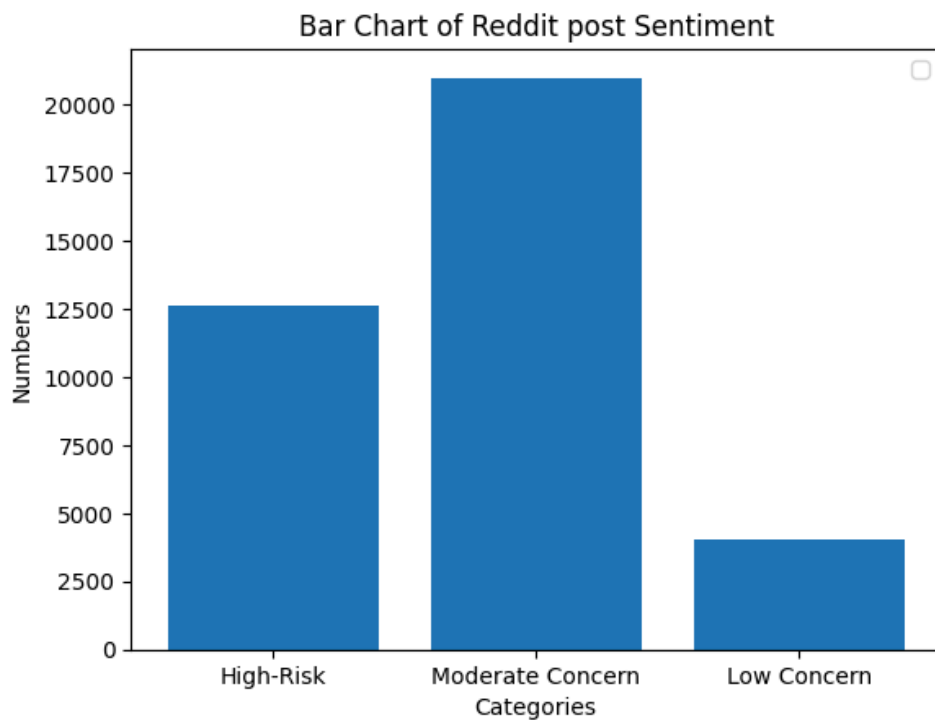


Fig 4. Bert trained model evaluation

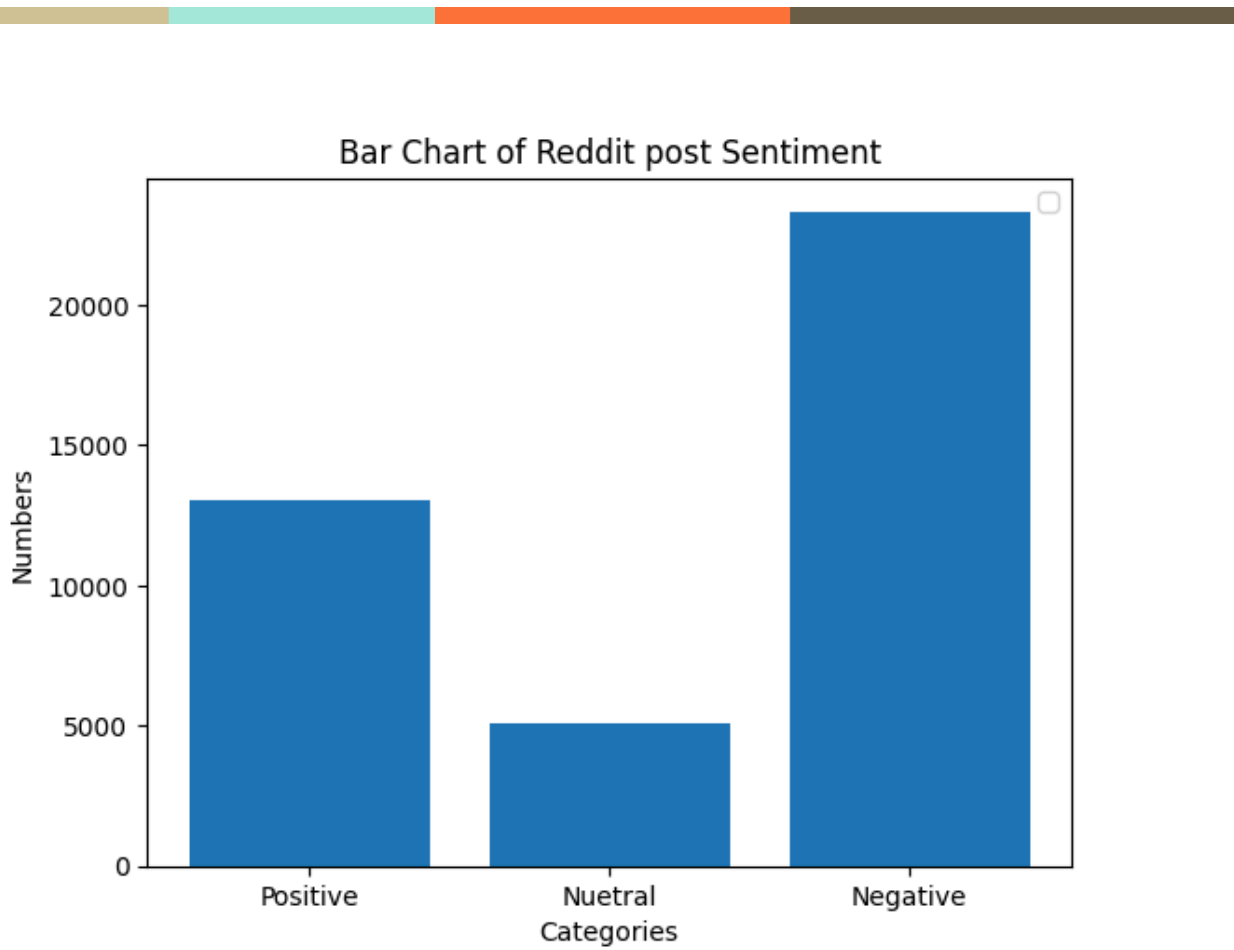


Fig 2. Vader analysis on Reddit scraped data

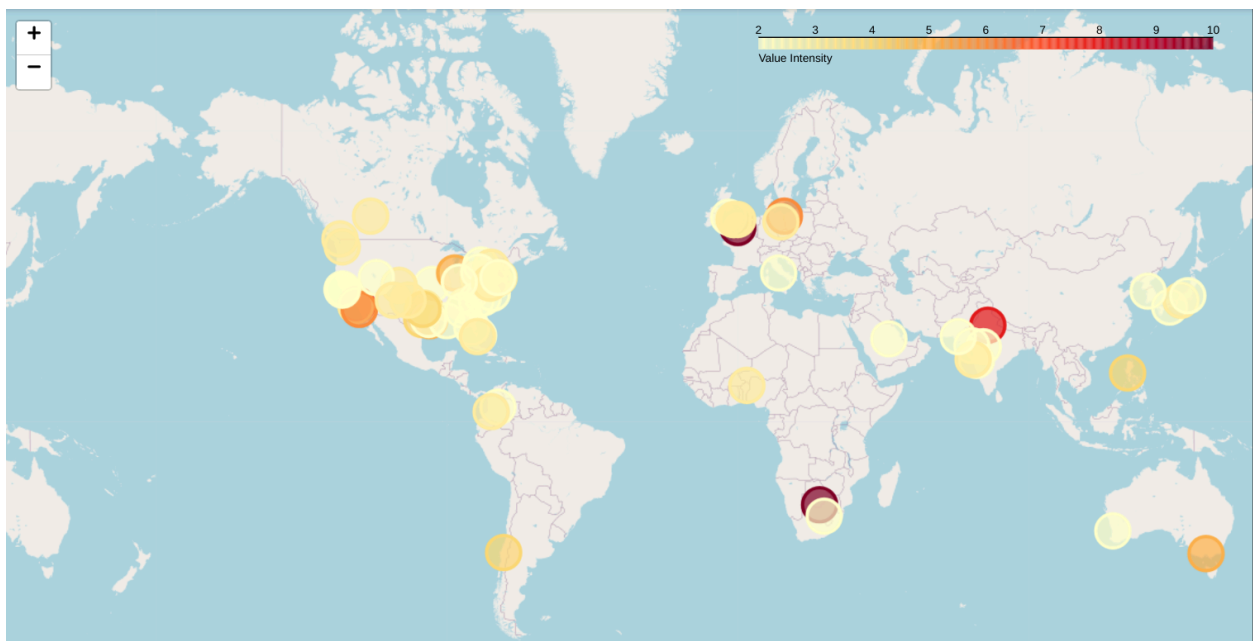


Fig 5. Folium map with color proportional to victim concentration

