Unveiling Mental Health Insights through Social Media Analysis

DATS 6312 – Natural Language Processing

Anjali Mudgal

Table of Contents

Introduction

The prevalence of mental health issues globally is staggering, with nearly 1 billion affected individuals—over 10% of the world's population—according to the World Health Organization. In the US alone, more than 57.8 million adults, constituting over 20% of the population, reported living with mental health conditions in 2021. This rise in mental health concerns parallels a significant shift towards social media as the primary means of communication. These platforms not only connect us but also offer insights into mental health indicators, making the connection between mental health and social media an increasingly critical area for exploration.

2. Dataset

2.1 Dataset Used

Through a meticulous literature review, we identified a research paper (https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10315126) closely aligned with our research goals. Leveraging the dataset from this paper, users' English and Spanish tweets were collected via the Twitter API, focusing on individuals openly sharing their diagnoses from September 1st, 2020, to August 31st, 2021. Retrieving up to 3200 of the most recent tweets per user, the dataset underwent refinement, excluding retweets and non-target language content, within a 5-year posting period. To ensure a comparative basis, a control group with similar tweet counts and posting periods was established. This rigorous data collection and filtering process underpin our investigation, ensuring a comprehensive and focused analysis.

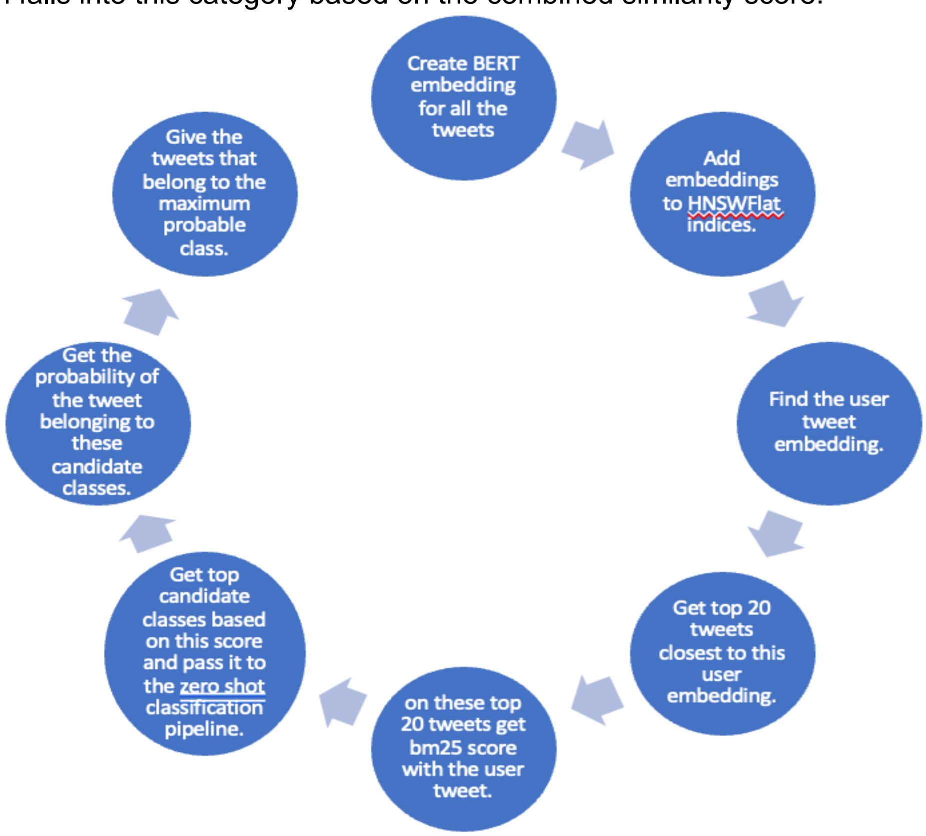
The tweets were classified into 10 different classes based on the mental health disorder. The classes were: EATING DISORDER, SCHIZOPHRENIA, OCD, PTSD, ANXIETY, BIPOLAR, AUTISM, DEPRESSION, ADHD, CONTROL.
Control Group was the non-diagnosed group of users.

3. MY CONTRIBUTION

Added python scripts to sample and divide equal number samples to train, val, test, dev in each of our instances.

Similarity scores and giving similar tweets if detected with a mental disorder.
We calculated the scores as follows:

1. Created BERT embeddings from the model that we fine-tuned using BertForSequenceClassification.
2. For all the tweets by removing the classification layer and getting the output bert embedding of CLS for each tweet.
3. These embeddings were used to create HNSWFlat indexes from FAISS. Later, the embeddings were added to the clusters.
4. Whenever the user enters a tweet, we calculate the BERT embedding for that user and search for the top 20 closest tweets in the HNSW. (It returns the distance, which we are using to get the similarity)
5. Using these tweets we are finding the Okapi BM25 score for each tweet with the query tweet.
6. The combined score of each embedding with these 20 tweets would be 0.8*(similarity from HNSW bert embedding matching) + 0.2*(bm25 similarity)
7. We now get the top similar tweets, we would use the unique classes in these tweets to do a zero shot classification to get the maximum probable class in which the tweet can fall into.
8. After getting the top class, we are displaying the tweets in the top 20 selection from hnsw which falls into this category based on the combined similarity score.



4. Conclusion

We learned that passing keyword information can help with text analysis where types of words, punctuations used also matter.

We also understood the vector based databases and how HNSW or similar vector based similarity algorithms can help improve the real time search results.

5. Further Improvements

- The development of a mental health assistance chatbot to provide necessary support for the user proved to be challenging for us. However, we plan to try to incorporate this in the future.

- We can also try implementing the langdetect model to detect the language in which a tweet was posted and classify it accordingly.

- We could try langdetect model to detect the language in which the tweet was posted and then use the classification model accordingly.

Percentage of code copied from internet: Around 25%

6. References:

1. https://huggingface.co/dccuchile/distilbert-base-spanish-uncased-finetuned-xnli/tree/main
2. https://huggingface.co/bert-base-uncased
3. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10315126

4. https://www.researchgate.net/publication/361757484_Mental_Illness_Classification_on_Social_Media_Texts_using_Deep_Learning_and_Transfer_Learning
5. https://nayakpplaban.medium.com/mental-health-classification-using-distilbert-and-supervised-machine-learning-d4b0c536f7b2
6. https://medium.com/@zz1409/combining-embedding-and-keyword-based-search-for-improved-performance-b15b0cfd3152
7. https://www.pinecone.io/learn/series/faiss/faiss-tutorial/
8. https://www.pinecone.io/learn/series/faiss/vector-indexes/