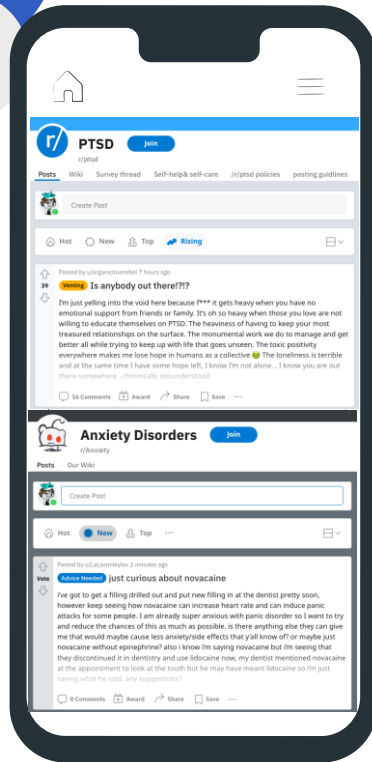




Distinguishing **PTSD** from **Anxiety Disorder**: A Machine Learning Investigation of Linguistic Patterns in Online Mental Health Communities

Youngmeen Kim

Department of Applied Linguistics
Georgia State University
ykim140@gsu.edu



Overlapping Symptoms

PTSD - Recurrent emotional responses to trauma or frightening memories (Brewin et al., 2000)

Anxiety Disorders – Persistent anxiousness or fear, which can take many forms (American Psychiatric Association, 2013)

PTSD Symptoms



[verywellhealth.com/what-is-ptsd-5084527](https://www.verywellhealth.com/what-is-ptsd-5084527)

Signs of an Anxiety Disorder



[verywellhealth.com/do-i-have-anxiety-5207282](https://www.verywellhealth.com/do-i-have-anxiety-5207282)

People turn to Online Communities

- Only 47.2% of individuals with mental illness in the US receive professional treatment (SAMHSA, 2021).
- Individuals with PTSD and anxiety disorders often experience **stigma** and may face **barriers** to accessing health care (e.g., Alonso et al., 2018; Sareen et al., 2016).
- Through online communities, individuals with these disorders connect with others who share **similar experiences** and access **resources** and **support** that may not be readily available to them (e.g., Moorhead et al., 2013; Naslund et al., 2016).
- Need to have a better **understanding of the individuals in online mental health** communities.

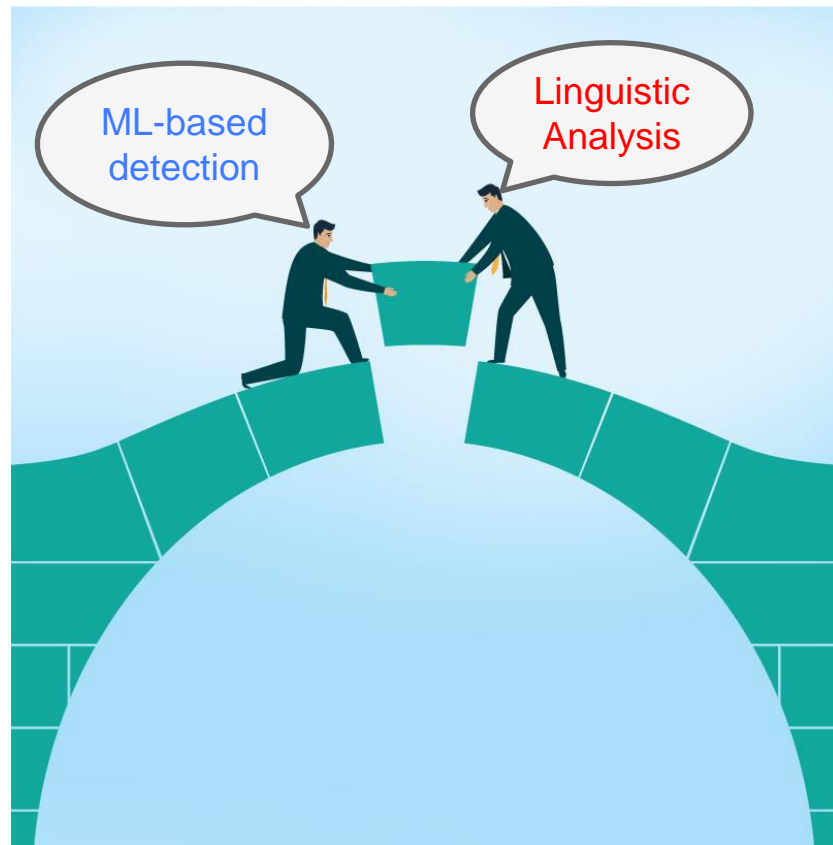


Bridging the Gap

With the development of NLP and ML, scholars in various fields have significantly improved mental health text detection models.

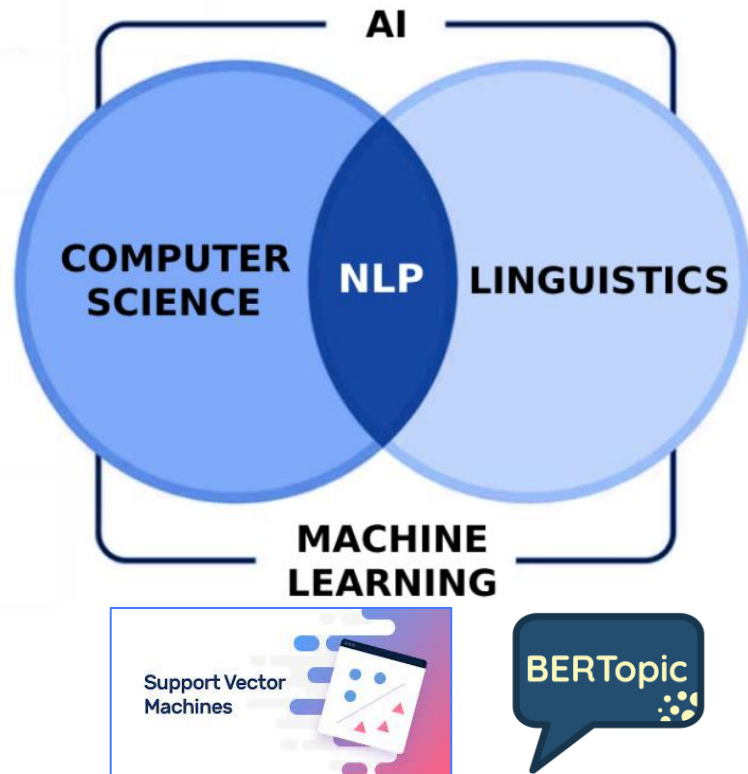
Previous studies using ML approaches significantly increased the accuracy of such detection models.

- But: Many **have not interpreted** the unique **linguistic features** in these types of texts in-depth.
- Many studies used **neutral** texts (e.g., daily life, travel) **as a comparison group** with mental health texts, which might not accurately reveal their distinctive and unique features.
- The **major topics** of the texts from the online mental health community have **not been interpreted** in accordance with findings from ML results.



Research Questions

- Can an **ML model accurately classify** texts from PTSD and anxiety disorder online communities?
- What are the **major topics** within the PTSD and anxiety disorder online communities?
- What are typical **language patterns** in texts related to PTSD and anxiety disorder?



Data

PTSD (Group1)

$N = 25,000$ posts
Total Number of Tokens =
approximately 4,027,500

About Community

We are a supportive, respectful community for discussion and links of interest for people who have PTSD or have friends, family members, or partners with PTSD. NOTE: We can not diagnose you with PTSD here. If you think you have PTSD, please see a licensed professional.

Created Sep 3, 2008

88.0k Members
148 Online

Create Post

Both Subreddits

Share experiences

Seek **support** from others

Similar **themes** and **topics**
(personal challenges, struggles,
seeking advice and information,
coping strategies)



Anxiety (Group2)

$N = 25,000$ posts
Total Number of Tokens =
approximately 3,725,000

About Community

Discussion and support for sufferers and loved ones of any anxiety disorder.

Created Sep 15, 2008

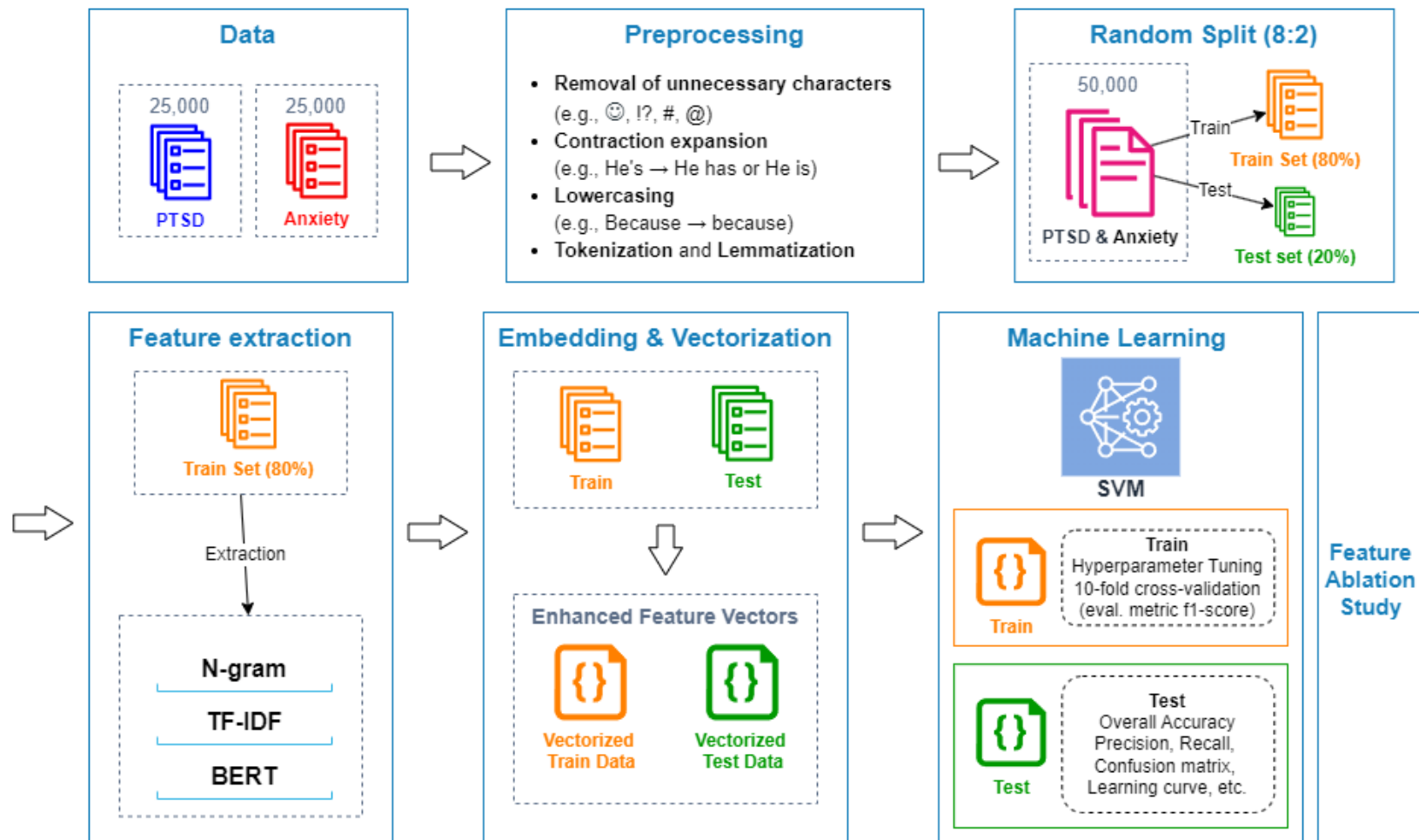
605k Members
1.0k Online

Create Post

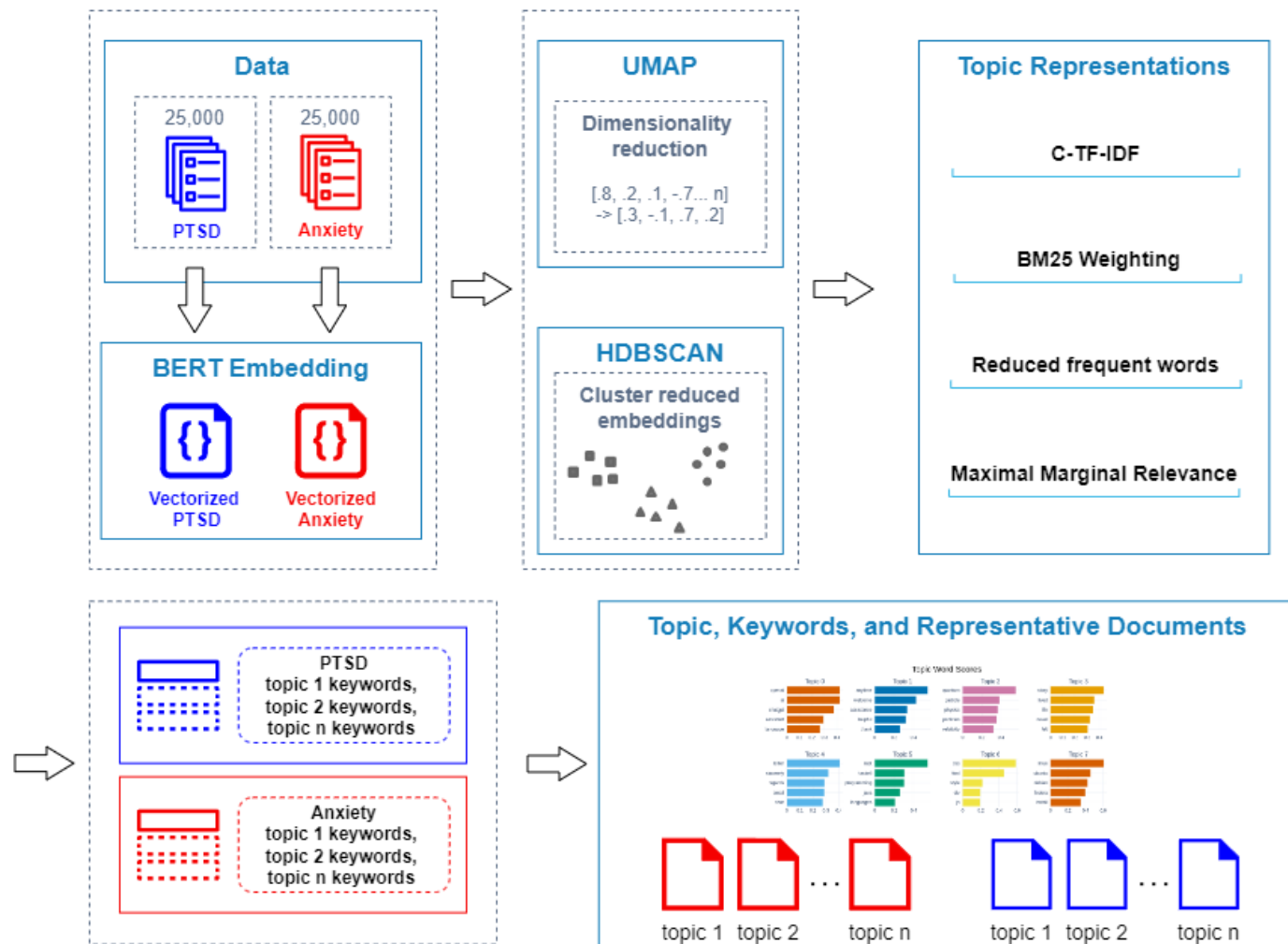
COMMUNITY OPTIONS

Community theme

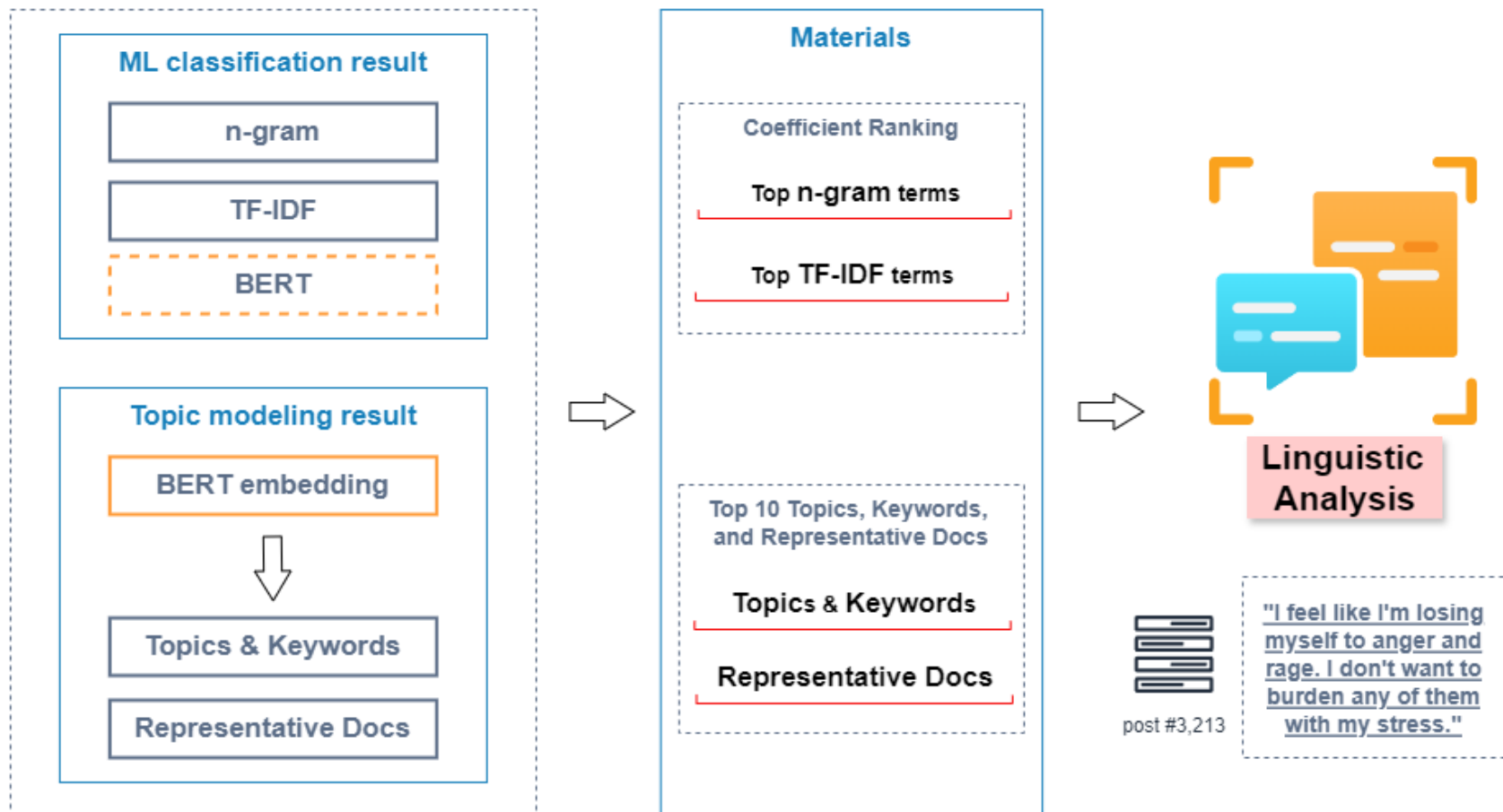
Machine Learning Workflow



BERTopic



Linguistic Analysis



Results

N-gram, TF-IDF, BERT

Classification

Classification Accuracy **88.93%**

Table 1.

Classification Result

| Type | precision | recall | f1-score | support |
|---------|-----------|--------|----------|---------|
| PTSD | 0.90 | 0.88 | 0.89 | 5,016 |
| Anxiety | 0.88 | 0.90 | 0.89 | 4,984 |

Table 2.

Confusion Matrix

| | | Actual | | |
|-----------|---------|--------|---------|--------|
| | | PTSD | Anxiety | Total |
| Predicted | PTSD | 4,403 | 613 | 5,016 |
| | Anxiety | 494 | 4,490 | 4984 |
| | Total | 4,897 | 5,103 | 10,000 |

Results

N-gram, TF-IDF (Without BERT)

Classification

Classification Accuracy **88.78%**

Table 3.

Classification Result

| Type | precision | recall | f1-score | support |
|---------|-----------|--------|----------|---------|
| PTSD | 0.90 | 0.87 | 0.89 | 5,016 |
| Anxiety | 0.88 | 0.90 | 0.89 | 4,984 |

Table 4.

Confusion Matrix

| | | Actual | | |
|-----------|---------|--------|---------|--------|
| | | PTSD | Anxiety | Total |
| Predicted | PTSD | 4,376 | 640 | 5,016 |
| | Anxiety | 482 | 4,502 | 4,984 |
| | Total | 4,858 | 5,142 | 10,000 |

Results

BERT Alone

Classification

Classification Accuracy **86.07%**

Table 5.

Classification Result

| Type | precision | recall | f1-score | support |
|---------|-----------|--------|----------|---------|
| PTSD | 0.87 | 0.84 | 0.86 | 5,016 |
| Anxiety | 0.85 | 0.88 | 0.86 | 4,984 |

Table 6.

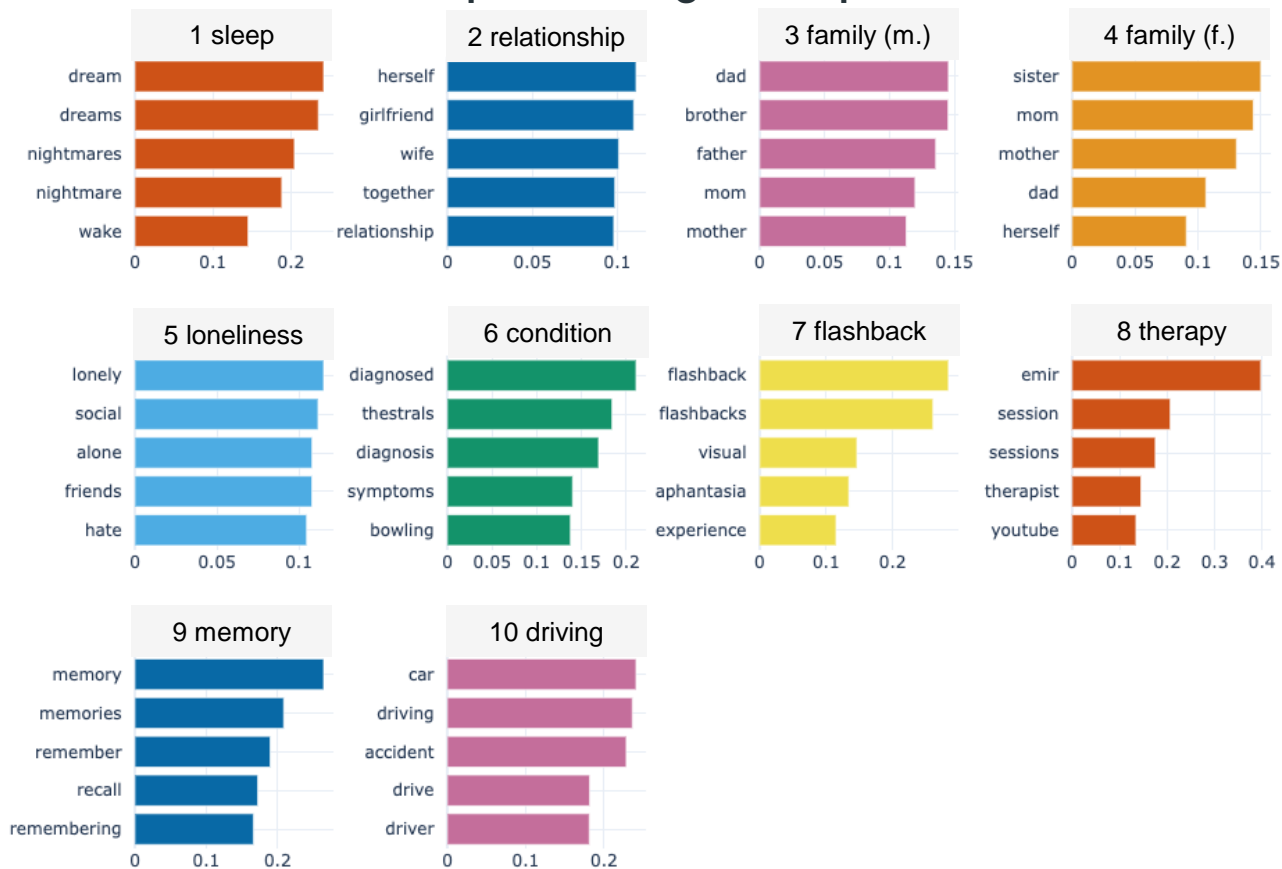
Confusion Matrix

| | | Actual | | |
|-----------|---------|--------|---------|--------|
| | | PTSD | Anxiety | Total |
| Predicted | PTSD | 4,234 | 782 | 5,016 |
| | Anxiety | 611 | 4,373 | 4,984 |
| | Total | 4,845 | 5,155 | 10,000 |

Results

Top Topics (PTSD)

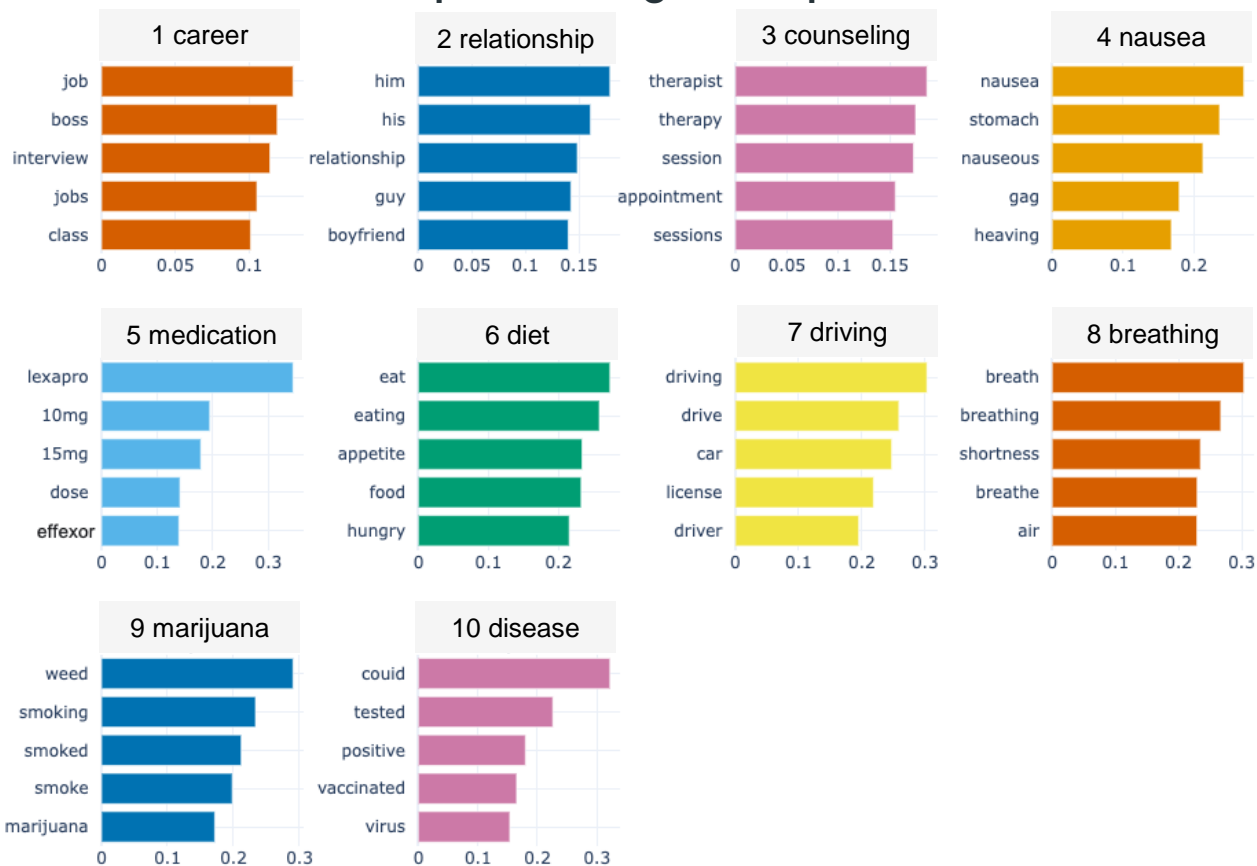
Topic Modeling (BERTopic)



Results

Top Topics (Anxiety)

Topic Modeling (BERTopic)



Result

Linguistic Features (PTSD)

Top N-grams & TF-IDF sorted by coefficient

Bigram

1. do you
2. **I still**
3. in a
4. have any
5. **talk about**
6. need to
7. It do
8. **not feel**
9. time i
10. he be
11. this be
12. be have
13. the first
14. could not
15. like a

Trigram

1. I have PTDS
2. diagnose with PTDS
3. **when I be**
4. **do not feel**
5. I be have
6. know it be
7. **if this be**
8. I be wonder
9. like it be
10. I do not
11. be diagnose with
12. I be diagnose
13. do not like
14. anyone have any
15. get out of

Quadgram

1. be diagnose with PTSD
2. **I can not get**
3. **and I can not**
4. be wonder if anyone
5. **I do not like**
6. do anyone have any
7. **I do not feel**
8. **do not even know**
9. I be tired of
10. but I feel like
11. I do not want
12. because I do not
13. I be so tired
14. it be go to
15. do not feel like

Quintgram

1. I be diagnose with PTSD
2. **I be so tired of**
3. **I do not even know**
4. **I do not want to**
5. **I do not feel like**
6. **do not know if this**
7. **because I do not want**
8. and I do not know
9. have be diagnose with PTDS
10. it be go to be
11. know what to do anymore
12. be not sure if this
13. I be wonder if anyone
14. I feel like I have
15. be not sure if this

TF-IDF

1. **anger**
2. do
3. one
4. **up**
5. would
6. fucking
7. happen
8. **physical**
9. **tell**
10. just
11. too
12. **someone**
13. **moment**
14. right
15. **pity**

Result

Linguistic Features (PTSD)

From N-gram & TF-IDF

1. **I'm diagnosed with PTSD** and still struggling to recognize my **symptoms**.
2. **I'm so tired of** avoiding telling people about my **trauma** because I'm scared of making people uncomfortable.
3. **I don't even know** why I'm posting this here, I know no advice can **fix** me, [...] until I get those nights where my body just shuts down and sleeps long enough
4. I have only two or three close people in my life but **I don't want to burden** any of them with my issues.
5. I know I'm an adult now but **I don't feel like** an adult, **I feel like** a child all the time and I don't know what I'm doing.

1. I feel like I'm losing myself to **anger** and **rage**.
2. I have no idea if the **abuse** was **physical**, sexual, verbal, or all of the above.
3. A few years ago I witnessed **someone** **commit** by jumping off a building.

5-grams

1. *I be diagnose with PTSD*
2. *I be so tired of*
3. *I do not even know*
4. *I do not want to*
5. *I do not feel like*

TF-IDF

1. *anger*
2. *physical*
3. *someone*

Results

Linguistic Features (Anxiety)

Top N-grams & TF-IDF sorted by coefficient

| Bigram | Trigram | Quadgram | Quintgram | TF-IDF |
|-------------------------|--------------------------|------------------------------|-----------------------------------|-----------------------|
| 1. <i>my anxiety</i> | 1. <i>now I be</i> | 1. <i>I can not help</i> | 1. <i>but I can not help</i> | 1. <i>plan</i> |
| 2. <i>the past</i> | 2. <i>I be scared</i> | 2. <i>I can not stop</i> | 2. <i>I do not know why</i> | 2. <i>only</i> |
| 3. <i>and I</i> | 3. <i>I be a</i> | 3. <i>and now I be</i> | 3. <i>I do not know how</i> | 3. <i>picture</i> |
| 4. <i>go to</i> | 4. <i>to the point</i> | 4. <i>most of the time</i> | 4. <i>I be not able to</i> | 4. <i>myself</i> |
| 5. <i>do I</i> | 5. <i>a panic attack</i> | 5. <i>I be so scared</i> | 5. <i>feel like I do not</i> | 5. <i>physically</i> |
| 6. <i>make I</i> | 6. <i>be go to</i> | 6. <i>I do not have</i> | 6. <i>for as long as I</i> | 6. <i>place</i> |
| 7. <i>now I</i> | 7. <i>my anxiety be</i> | 7. <i>to the point where</i> | 7. <i>really do not want to</i> | 7. <i>platform</i> |
| 8. <i>have a</i> | 8. <i>and I have</i> | 8. <i>but I can not</i> | 8. <i>how do you deal with</i> | 8. <i>placebo</i> |
| 9. <i>and have</i> | 9. <i>but I just</i> | 9. <i>I be go to</i> | 9. <i>do not know why I</i> | 9. <i>photograph</i> |
| 10. <i>since I</i> | 10. <i>go to the</i> | 10. <i>and I have to</i> | 10. <i>do knot know how I</i> | 10. <i>internally</i> |
| 11. <i>I be</i> | 11. <i>and I be</i> | 11. <i>I just can not</i> | 11. <i>I can not stop think</i> | 11. <i>physician</i> |
| 12. <i>know how</i> | 12. <i>I go to</i> | 12. <i>I really do not</i> | 12. <i>I feel like I be</i> | 12. <i>pic</i> |
| 13. <i>like I</i> | 13. <i>like I be</i> | 13. <i>I have be feel</i> | 13. <i>long as I can remember</i> | 13. <i>sometimes</i> |
| 14. <i>panic attack</i> | 14. <i>and now I</i> | 14. <i>I be afraid of</i> | 14. <i>I be do to die</i> | 14. <i>placid</i> |
| 15. <i>be in</i> | 15. <i>but I can</i> | 15. <i>like I can not</i> | 15. <i>just do not want to</i> | 15. <i>mad</i> |

Results

Linguistic Features (Anxiety)

From N-grams & TF-IDF

1. thinking I might die or am seriously ill with **no real reason to think** so. I know I am overreacting **but I can't help** it
2. So I said sorry to her and said **I don't know why** I am feeling **uncomfortable**.
3. **I don't know how** much longer I can handle **living life** like this and not being able to live life the way I want to because of my anxiety.
4. I'm so sad 'cause **I'm not able to** control my **fears**, I cannot stop thinkin about
5. I don't know but I really miss that tiny moment of calmness. I **feel like I don't** get any calmness anymore

1. I feel like I'm going to **pass out** when I think about going cause I **picture** the doctor saying I only have a few months left to live.
2. Even if the thing I am **worrying** over is something I've told **myself** is **not logically plausible**, I still feel incredibly anxious and like I'm about to **die**.
3. I stayed calm in the outside for my kids but **internally** I'm **struggling**.

5-grams

1. *but I can not help*
2. *I do not know why*
3. *I do not know how*
4. *I be not able to*
5. *feel like I do not*

TF-IDF

1. picture
2. myself
3. internally

Implications

The model can help build a more accurate and practical PTSD and Anxiety **detection/classification** system.

The results contribute to a better understanding of **anonymity** practices online around mental health concerns.

The unique **linguistic characteristics** and the topics of PTSD and Anxiety texts could be helpful for various **practical applications** (e.g., diagnosis, treatment).



Thank you for your attention.

Youngmeen Kim

Department of Applied Linguistics
Georgia State University

ykim140@gsu.edu



EMORY
UNIVERSITY

GeorgiaStateUniversity® 