Comparing panic and anxiety on a dataset collected from social media

Sandra Mitrović[1]
Oscar William Lithgow-Serrano[1]
Carlo Schillaci[2]
[1]IDSIA - USI/SUPSI, Switzerland
[2]SUPSI, Switzerland
sandra.mitrovic@idsia.ch, oscarwilliam.lithgow@idsia.ch
carlo.schillaci@student.supsi.ch

Abstract

The recognition of mental health's crucial significance has led to a growing interest in utilizing social media text data in current research trends. However, there remains a significant gap in the study of panic and anxiety on these platforms, despite their high prevalence and severe impact. In this paper, we address this gap by presenting a dataset consisting of 1,930 user posts from Quora and Reddit specifically focusing on panic and anxiety. Through a combination of lexical analysis, emotion detection, and writer attitude assessment, we explore the unique characteristics of each condition. To gain deeper insights, we employ a mental health-specific transformer model and a large language model for qualitative analysis. Our findings not only contribute to the understanding digital discourse on anxiety and panic but also provide valuable resources for the broader research community. We make our dataset, methodologies, and code available to advance understanding and facilitate future studies.

1 Introduction

The indisputable importance of mental health is recently reflecting in the increased research interest, putting the emphasis on identifying related issues in the textual sources in social media and focusing mainly on depression (William and Suhartono, 2021; Bhadra and Kumar, 2022; Parapar et al., 2023) and suicide (Bayram and Benhiba, 2022; Malhotra and Jindal, 2022). We contend that anxiety and panic, despite being understudied, are equally significant. Panic is conventionally described as a sudden, overwhelming fear (Bloom et al., 2009), while anxiety is marked by persistent unease and uncontrollable worry (Stein and Sareen, 2015). Despite their shared characteristics, the literature indicates that the differentiation between generalized anxiety and panic is valid (Russell Noyes et al., 1992). While anxiety, in the related literature, is typically coupled with depression (Burkhardt et al., 2022; Tasnim et al., 2023), panic is, in general and with few exceptions (Mitrović and Kanjirangat, 2022), (Mitrović et al., 2023), far less studied. Discriminating between anxiety and panic is important, given that individuals experiencing panic face a higher risk of more profound psychological and psychiatric issues, including acute suicidality and agoraphobia. Moreover, the study of panic can be instrumental in identifying individuals with PTSD, making it a highly worthwhile topic. Thus, the ability to identify patients undergoing panic is crucial, carrying substantial clini-

ANXIETY: "I honestly don't think it can be even described in words. I think anxiety comes in different ways in different people, so i am just going to say about my experience. Having anxiety is not being able to remember the last time that you were relaxed without a disturbing thought in your head. Having anxiety means that your brain is extremely creative with coming up with the worst improbable almost impossible scenario. It is like having a brain that thinks every minor chest pain as a heart attack, every headache a brain tumor and every numbness ms. Having anxiety means to always overthink every action and saying and every situation. It means that you have to forget about having fun and enjoying life. It means that you have always to doubt yourself. I can literally go on forever."
PANIC: "Picture yourself sitting on the couch enjoying a television show. You feel relaxed and decide to get up and get something to drink. You open the cabinet door to get a glass when... our of nowhere you feel this big surge of suffocating fear.. your heart starts pounding out of your chest, reality suddenly doesn't feel real. All you know is that you must be alone. You feel like you are free falling faster and faster into the dark pit of hell. You feel depressed and defeated and when it passes you feel exhausted."

Figure 1: An example of an ANXIETY and a PANIC post extracted from Quora website.

Source	Panic	Anxiety	*Total*
Quora	526	976	1502
Reddit	187	241	428
Total	**713**	**1217**	**1930**

Table 1: Number of samples per source and per class.

Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2024), pages 153–165
March 21, 2024

cal implications. Nevertheless, to the best of our knowledge, current literature lacks computational approaches for discerning panic and anxiety in social media textual resources. To bridge this gap, we collected a related dataset of 1,930 user posts originating from two well-known websites, Quora[1] and Reddit[2], and we conduct a versatile analysis to address the following research questions:

[RQ1]: How (dis)similar are panic-annotated to the anxiety-annotated posts? And how different are the posts coming from different blogs?

[RQ2]: How successfully can a classifier discern a panic post from an anxiety one?

[RQ3]: What a qualitative NLP-assisted analysis on this dataset can tell about mental health and related context?

Besides the insights of various lexical, emotion, writers' attitudes, classical machine learning and (large) language models-based approaches, we contribute the literature by providing the dataset, analysis and code[3].

We anticipate that certain discoveries from our research may contribute to practical applications. In particular, distinguishing between anxiety and panic triggers holds potential clinical utility and could guide the deployment of emergency medical assistance. Additionally, it may encourage individuals to reach out to their designated support person. In general, the paper could aid in screening posts on social media displaying indications of anxiety or panic, thereby contributing to a better mental health understanding and practice.

2 Dataset collection

We web scraped the data[4], starting from the well-known question answering website Quora and a set of questions related primarily to panic (attacks), extending it to similar questions related to anxiety (or other anxiety-related questions that we came upon in Quora). We then switched to another popular blogging website, Reddit, and looked for semantically similar questions and answers. We only collected the original question and the first reply to it, without tracking the whole conversation (replying to a reply is frequently common in Reddit). Question could contain either "anxiety" or "panic" keyword and we applied rule-based annotation based on the keyword presence (considering panic class as positive). It is important noting that questions, albeit used for determining class label, are not considered as the integral part of the dataset.

Tab. 1 presents basic distribution of posts per class per blog, while the questions used for data collection can be found in Appendix A.2, Tab. 4. It is noteworthy to mention that we adhered to this predetermined set of questions to guarantee the distinct separation of classes, guided by the rule-based strategy driven by question formulation.

3 [RQ1] How (dis)similar are panic- to the anxiety-annotated posts? And is there any difference wrt the post source (originating blog)?

To this end, we conducted a multifaceted analysis.

Linguistic Perspective Using LIWC-22 (Pennebaker et al., 2022) features, we could see than panic posts have on average more words than those of anxiety (169.87 vs. 160.83), also containing more pronouns and verbs. Anxiety posts, on the other hand, contain on average more words per sentence (20.64 vs 18.47) and longer words (23.17 vs. 17.45) than panic (see A.2, Fig. 5). Quora posts contain more words in general (191.64 vs. 67.77), more words per sentence and longer words than Reddit ones, which, on the contrary, are richer in verbs, adverbs and pronouns (see A.2, Fig. 6).

Readability Perspective Upon calculating different readability scores we have noticed prominent difference between two classes on average Flesch score (FS) for Quora posts, with 49.84 for anxiety versus 63.72 for panic, meaning that in Quora, posts denoted as anxiety are more difficult to read than those denoted as panic (see A.2, Fig. 9). Reddit texts are in general, more similar with respect to readability and also easier to read (with average FS per both classes being around 70).

Emotion Perspective We analyzed two different aspects related to emotions: their presence and their intensity. To determine how represented the emotions are in a post, we employed a pretrained HuggingFace language model (Demszky et al., 2020) and used inferred emotion probabilities relative to 28 different emotions, from the Go-Emotions dataset, as a proxy.

When considering only 5 most represented emotions per post (Fig. 2), we can notice that gratitude, nervousness and pride have higher average rank in the anxiety posts than in the panic ones. Similar

[1] https://www.quora.com/
[2] https://www.reddit.com/
[3] https://github.com/SandraMNE/QRPanicAnxiety
[4] For implementation details see Appendix A.1.

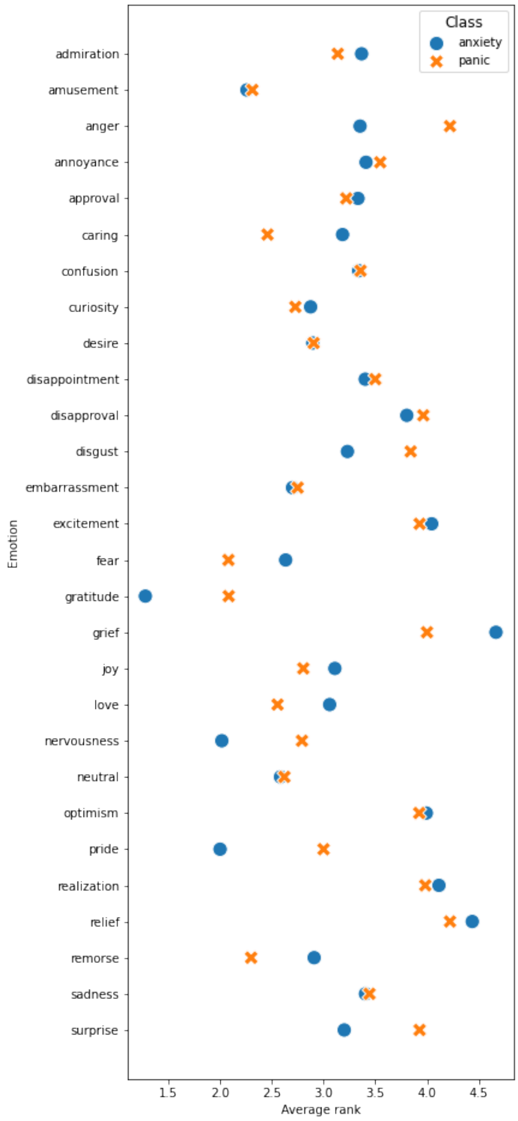

Figure 2: Average rank of emotions, calculated based on the 5 most important emotions per post.

pattern, although with much less important average rank, can be observed for disgust and surprise, while fear, remorse, caring and love have higher average rank in the panic posts. Within the top 5 most represented emotions per post, the 5 most frequent are neutral emotion, nervousness, fear and approval for both panic and anxiety; realization for panic and caring for anxiety (see A.2, Fig. 4).

When considering all the emotions per post, the Mann-Whitney U-test revealed a statistically significant difference in the mean ranks of fear, admiration, amusement, approval, disgust, gratitude, optimism, realization and surprise emotions between panic and anxiety posts, with a significance level set at $\alpha = 0.05$.

To assess emotion intensities, we used intensity lexicon (Mohammad, 2018) considering only 8 emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, trust. As expected, fear is the emotion with the highest intensity in both classes. In panic posts, fear and anger have higher intensity on average than in anxiety posts (fear intensity: 0.46 in anxiety vs. 0.53 in panic posts; anger intensity: 0.34 in anxiety vs. 0.40 in panic posts). Conversely, anticipation and sadness have higher average intensity in anxiety posts than in the panic ones (see Fig. 7). With regard to the post sources, as displayed in A.2, Fig. 8, all considered emotion intensities are amplified in Quora as compared to Reddit.

Writer's Attitude Perspective We investigate writer's attitude from the perspective of convincingness (Gretz et al., 2019), persuasiveness and usage of irony (Barbieri et al., 2020). While almost no difference could be noted in the irony and persuasiveness average scores between the two classes, average score for convincingness somewhat differs (0.76 for anxiety vs. 0.73 for panic). When considering the sources, we observe that panic posts in both sources have similar convincingness level. However, on overall Quora posts exhibit higher average convincingness (0.78 vs. 0.7) and persuasiveness scores (0.36 vs. 0.29) than Reddit. On the contrary, Reddit posts contain more irony (average score 0.17) than Quora's (0.12).

4 [RQ2] How successfully can a classifier discern panic from anxiety posts?

In order to answer this question, we build a Gradient Boosting classifier. Apart from already mentioned LIWC variables (denoted as L), emotion probabilities (denoted as M), emotion intensities (denoted as I), convincingness, persuasiveness and irony scores (denoted together as W), we calculated also the embeddings (B) of the posts and included them as input features. The obtained results can be seen in Tab. 2. Fairly good performance of classifier, low variance across 10 runs and relatively small discrepancy between F1-macro and F1-weighted/F1-micro scores, despite the strong class imbalanceness, showcase the good predictive power of features and classifier robustness. Additionally, sacrificing a bit of performance by excluding embeddings we can obtain an explainable model. The latter, as depicted in A.2, Fig. 10, showcases that LIWC mental, fear intensity and LIWC long words are considered important in all

10 runs.

Features	F1-weighted	F1-macro	F1-micro	ROC AUC	MCC
B	0.896 (0.01)	0.889 (0.01)	0.897 (0.01)	0.960 (0.01)	0.780 (0.02)
L+M+I+W	0.816 (0.02)	0.801 (0.02)	0.819 (0.02)	0.891 (0.02)	0.609 (0.03)
B+L+M+I+W	0.899 (0.01)	0.891 (0.01)	0.898 (0.01)	0.960 (0.01)	**0.783 (0.02)**

Table 2: XGBoost classifier average performance in terms of F1-score, ROC AUC and MCC (and standard deviation) across 10 runs. Notation: B - embedding, L - LIWC, M - emotion probabilities, I - emotion intensities, W - writer's attitude.

Probing the rule-based annotation One might argue that the rule-based class separation used for annotation makes the two classes easily distinguishable based on the two keywords. We therefore perform two straightforward yet efficient analysis.

First, we investigated whether the imposed one-question-one-keyword pattern holds for the answers as well. It turned out that it was not the case. More specifically, out of total 1930 blog posts, as much as 386 (20%) contain both "panic"and "anxiety" keywords. The co-occurrence of "panic"and "anxiety" keywords is specifically prominent in Quora where 343 out of 1502 posts (22.84%) contain both keywords, while in Reddit this is less frequent (only 43 out of 428 posts or 14.06%).

Second, we replaced all the occurrences of the two keywords with "MASKEDCONTENT" token, recalculated the respective features on the obtained texts and retrained the classifier. As could be seen from Tab. 3, although, as expected, the performances drop in terms of all considered metrics, the downgrade in the performance is not prominent. This indicates that the presence of the "panic" and "anxiety" keywords in not crucial for classification.

5 [RQ3] What a qualitative NLP-assisted analysis on this dataset can tell about mental health and related context?

We conduct a two-fold analysis. First, we aim to determine what the posters were experiencing. To this end, we prompted language model pretrained on mental health discussions (Ji et al., 2021)), requesting a YES/NO answer related to stress and open answers related to "experiencing". Results are depicted in A.2, Fig. 12 and 11 respectively.

Second, we leveraged Large Language Models (LLM), specifically ChatGPT-3.5. We initially used it to assess the dataset annotation quality, finding that only 8.4% of collected panic posts were misclassified as anxiety by ChatGPT, and merely 3.8% of original anxiety posts were erroneously labeled as panic. We then conducted a comprehensive multidimensional thematic and content analysis. This involved automated coding and key-phrase extraction across various dimensions (see A.2 for details). Several noteworthy observations emerged from our investigation. Notably, panic attacks exhibited a higher "Strong" intensity (21%) compared to anxiety attacks (7%). In the context of relationships, the top four categories were shared among classes. However, the fifth differed, with "Work" for anxiety and "Romantic partners" for panic. Location-wise, the prevalent places were Home, School, Work, and Public places. In the triggers category, social events constituted 33% of anxiety triggers, contrasting with only 9% for panic attacks. Traumatic experiences accounted for 11% of panic triggers, as opposed to 5.6% for anxiety. Notably, financial stress was among the top five anxiety triggers but was absent in panic (see Fig. 3). Understanding anxiety and panic triggers is vital for uncovering their root causes and informing targeted interventions (Johnson et al., 2014; Craske, 1991; Barzegar et al., 2021). Examining specific triggering events helps elucidate factors contributing to distress. Focusing on locations where these experiences occur provides insights into environmental influences (Swee et al., 2021), aiding the development of strategies for supportive environments. Investigating relationships contributes valuable insights into the impact of social interactions on these conditions (Tonge et al., 2020), facilitating interventions to improve social relationships and support systems. Figure 3 emphasizes these dimensions to enhance comprehension.

In the category of other_medical_conditions, stress-related conditions were prevalent in both anxiety and panic. However, panic exhibited a unique pattern, with breath-related conditions like asthma, COPD, sinus infections, nasal congestion, and COVID-19 being the most frequently mentioned. More details on different distributions, including professional inteventions, is provided in A.2, Tab. 5.

Features	Dataset	F1-weighted	F1-macro	F1-micro	ROC AUC	MCC
B	Original	0.896 (0.01)	0.889 (0.01)	0.897 (0.01)	0.960 (0.01)	**0.780 (0.02)**
	Masked	0.875 (0.01)	0.866 (0.01)	0.876 (0.01)	0.940 (0.01)	0.734 (0.02)
B+L+M+W	Original	0.896 (0.01)	0.889 (0.01)	0.897 (0.01)	0.960 (0.01)	0.779 (0.02)
	Masked	0.873 (0.01)	0.864 (0.01)	0.874 (0.01)	0.940 (0.01)	0.730 (0.02)

Table 3: XGBoost classifier average performance (and standard deviation) in terms of F1-score, ROC AUC and MCC across 10 runs on the original and masked dataset. Notation: B - embedding.

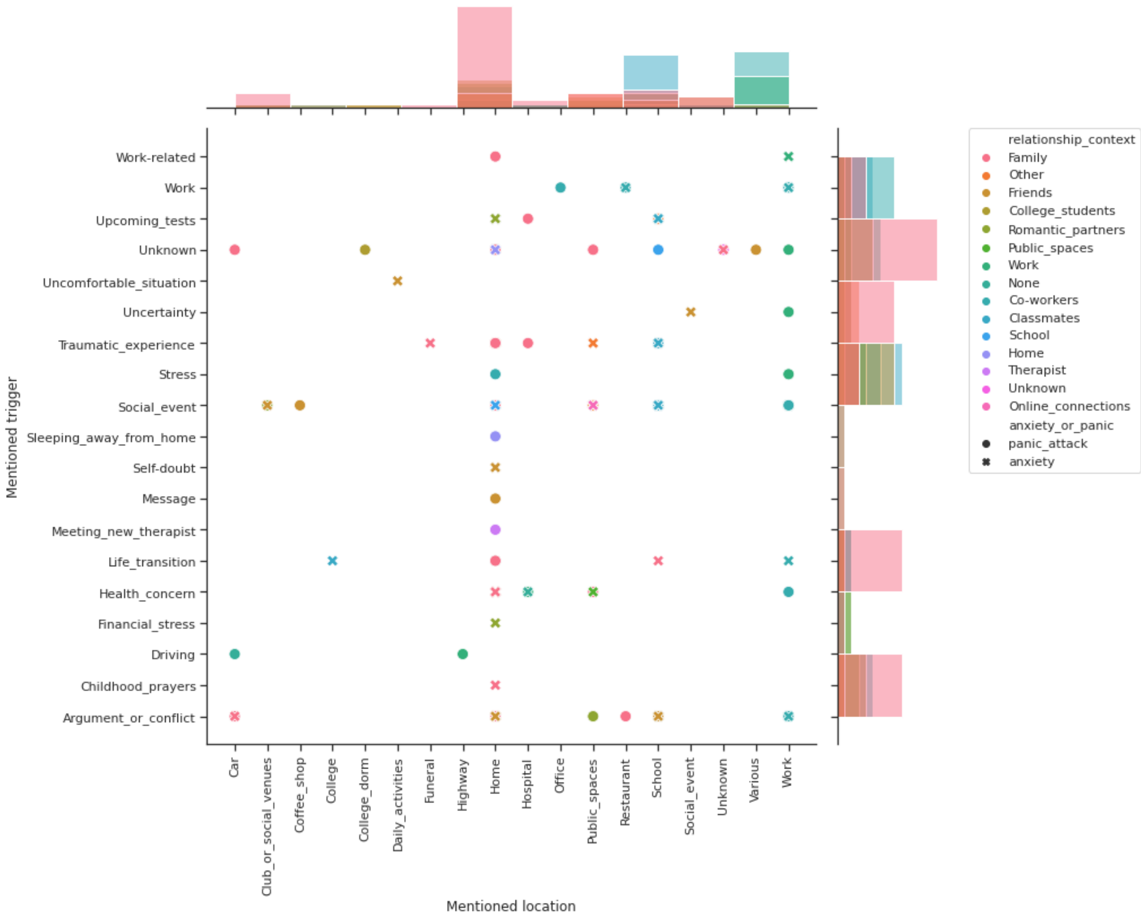

Figure 3: Context mentioned (Location, Social relationship and Trigger) in posts of personal experiences. Posts coded as "None" or "Other" on a specific axis were disregarded to enhance the readability.

6 Conclusion

In this study, we collected a dataset on two prevalent and closely related mental disorders: panic and anxiety. Our analytical exploration centered around three primary questions. Firstly, we focused into uncovering the distinctions and similarities, both in content and sources, between the anxiety and panic classes. To address this, we conducted lexical analyses, emotion assessments, and evaluations of writers' attitudes. Secondly, we assessed the efficacy of a classical machine learning classifier, leveraging combinations of the various extracted features, in distinguishing between posts belonging to these classes. Importantly, we conducted experiments to ensure the classifiers did not rely solely on the presence of explicit "panic"/"anxiety" keyword. Lastly, we sought insights into these disorders through a qualitative analysis of the dataset, leveraging a bidirectional transformer model trained on mental health texts (Mental-RoBERTa) and a large-language model (ChatGPT-3.5).

As a part of our contribution, we provide the entire output resulting from the NLP processes, including the collected codings and key-phrase extractions derived from the qualitative analysis. We hope that the obtained insights could be further exploited by domain experts. Looking ahead, future explorations aim to contribute to bring more layers of understanding within the context of these mental health disorders, for example, including a more in-depth analysis and clustering of key-phrase excerpts, with a particular focus on those associated with coping strategies, negative behavioral changes, and the impact on daily life.

7 Acknowledgements

The work of Oscar Lithgow-Serrano was partially supported by the project "Risk Identification and Prevention of Work-Related Stress Disorders" (Innosuisse 62423.1 IP-LS).

8 Limitations

As with any research, our study is not without limitations. With respect to data collection, the limitations are manifold. First, we considered limited number of questions which additionally have not been very well balanced between panic and anxiety. Second, the collected dataset is quite small and also the class distribution is quite skewed, hence we have to be very careful in assuming its generalizability. Third, as with all studies relying on the social media data, there is no possibility to know who is posting or how accurately. Fourth, we acknowledge that individuals writing may be experiencing anxiety, but they are unlikely to be blogging during a panic attack. Therefore, their narratives or teachings are more likely to revolve around recounting or educating about panic, rather than experiencing it in real-time.

Additionally, the emotion intensities were not calculated for all 28 emotions considered for the analysis of the emotions presence. Moreover, despite the small dataset size, we could have still tried to train a classifier with deep architecture.

9 Ethical considerations

Although Quora and Reddit data are publicly available, we do understand that the collected content is privacy-sensitive and that we might, even involuntarily, expose person's mental health-related privacy. To prevent easy backtrack to the author, we decided to delete not only users' usernames but also the original questions posed.

References

Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. 2020. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1644–1650, Online. Association for Computational Linguistics.

Habibeh Barzegar, Mostafa Farahbakhsh, Hosein Azizi, Sepideh Aliashrafi, Hossein Dadashzadeh, and Ali Fakhari. 2021. A descriptive study of agoraphobic situations and correlates on panic disorder. *Middle East Current Psychiatry*, 28(1):31.

Ulya Bayram and Lamia Benhiba. 2022. Emotionally-informed models for detecting moments of change and suicide risk levels in longitudinal social media data. In *Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology*, pages 219–225.

Sweta Bhadra and Chandan Jyoti Kumar. 2022. An insight into diagnosis of depression using machine learning techniques: a systematic review. *Current medical research and opinion*, 38(5):749–771.

Floyd E Bloom, Nicholas C Spitzer, Fred Gage, and Tom Albright. 2009. *Encyclopedia of Neuroscience, Volume 1*, volume 1. Academic Press.

Hannah Burkhardt, Michael Pullmann, Thomas Hull, Patricia Areán, and Trevor Cohen. 2022. Comparing emotion feature extraction approaches for predicting depression and anxiety. In *Proceedings of the eighth workshop on computational linguistics and clinical psychology*, pages 105–115.

Michelle G. Craske. 1991. Phobic fear and panic attacks: The same emotional states triggered by different cues? *Clinical Psychology Review*, 11(5):599–620.

Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. GoEmotions: A Dataset of Fine-Grained Emotions. pages 4040–4054.

Shai Gretz, Roni Friedman, Edo Cohen-Karlik, Assaf Toledo, Dan Lahav, Ranit Aharonov, and Noam Slonim. 2019. A Large-scale Dataset for Argument Quality Ranking: Construction and Analysis.

Shaoxiong Ji, Tianlin Zhang, Luna Ansari, Jie Fu, Prayag Tiwari, and Erik Cambria. 2021. MentalBERT: Publicly Available Pretrained Language Models for Mental Healthcare.

Philip L. Johnson, Lauren M. Federici, and Anantha Shekhar. 2014. Etiology, triggers and neurochemical circuits associated with unexpected, expected, and laboratory-induced panic attacks. *Neuroscience & Biobehavioral Reviews*, 46:429–454.

Anshu Malhotra and Rajni Jindal. 2022. Deep learning techniques for suicide and depression detection from online social media: A scoping review. *Applied Soft Computing*, page 109713.

Sandra Mitrović, Fabio Frisone, Suryam Gupta, Chiara Lucifora, Dragana Čarapić, Carlo Schillaci, Samuele Di Giovanni, and Ayushi Singh. 2023. Annotating panic in social media using active learning, transformers and domain knowledge. In *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 1269–1278. IEEE.

Sandra Mitrović and Vani Kanjirangat. 2022. Enhancing bert performance with contextual valence shifters for panic detection in covid-19 tweets. In *Proceedings of the 2022 6th International Conference on Natural Language Processing and Information Retrieval*, pages 89–92.

Saif M. Mohammad. 2018. Word affect intensities. In *Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*, Miyazaki, Japan.

Javier Parapar, Patricia Martín-Rodilla, David E Losada, and Fabio Crestani. 2023. erisk 2023: Depression, pathological gambling, and eating disorder challenges. In *European Conference on Information Retrieval*, pages 585–592. Springer.

JW Pennebaker, RL Boyd, RJ Booth, A Ashokkumar, and ME Francis. 2022. Linguistic inquiry and word count: Liwc-22. pennebaker conglomerates.

Jr Russell Noyes, Catherine Woodman, Michael J Garvey, Brian L Cook, Michael Suelzer, John Clancy, and Dorothy J Anderson. 1992. Generalized anxiety disorder vs. panic disorder: Distinguishing characteristics and patterns of comorbidity. *The Journal of nervous and mental disease*, 180(6):369–379.

Murray B Stein and Jitender Sareen. 2015. Generalized anxiety disorder. *New England Journal of Medicine*, 373(21):2059–2068.

Michaela B. Swee, Rachel M. Butler, Brennah V. Ross, Arielle Horenstein, Emily B. O'Day, and Richard G. Heimberg. 2021. Interpersonal Patterns in Social Anxiety Disorder: Predictors and Outcomes of Cognitive-Behavioral Therapy. *Cognitive Therapy and Research*, 45(4):614–627.

Mashrura Tasnim, Malikeh Ehghaghi, Brian Diep, and Jekaterina Novikova. 2023. Depac: a corpus for depression and anxiety detection from speech. *arXiv preprint arXiv:2306.12443*.

Natasha A. Tonge, Michelle H. Lim, Marilyn L. Piccirillo, Katya C. Fernandez, Julia K. Langer, and Thomas L. Rodebaugh. 2020. Interpersonal problems in social anxiety disorder across different relational contexts. *Journal of Anxiety Disorders*, 75:102275.

David William and Derwin Suhartono. 2021. Text-based depression detection on social media posts: A systematic literature review. *Procedia Computer Science*, 179:582–589.

A Appendix

A.1 Implementation details

Web scraping To collect the data, we employed Octoparse: https://www.octoparse.com/, a web scraping tool that is capable of recognising the structure of a web page and allows a user to select which specific content from the web page to extract.

Linguistic and psychometric features were obtained using LIWC-22, https://www.liwc.app/ (Pennebaker et al., 2022).

Readability scores were calculated using Python library: https://pypi.org/project/py-readability-metrics/.

Emotion probabilities were inferred using a pretrained HuggingFace language model SamLowe/roberta-base-go_emotions that can be found at: https://huggingface.co/SamLowe/roberta-base-go_emotions (Demszky et al., 2020). The list of 28 different emotions that are considered can be found at: https://github.com/google-research/google-research/blob/master/goemotions/data/emotions.txt.

Emotion intensities were calculated using intensity lexicon from: https://saifmohammad.com/WebPages/AffectIntensity.htm(Mohammad, 2018).

Writer's Attitude: **Convincingness** class softmax scores were obtained using fine-tuned HuggingFace model: https://huggingface.co/jakub014/bert-base-\uncased-IBM-argQ-30k-finetuned-\convincingness-acl2016 (Gretz et al., 2019). **Persuasiveness** class softmax scores were obtained using pretrained HuggingFace model: https://huggingface.co/paragon-analytics/roberta_persuade. **Irony** class softmax scores were obtained using HuggingFace model specifically fine-tuned for irony detection: https://huggingface.co/cardiffnlp/twitter-roberta-base-irony (Barbieri et al., 2020).

Classifier was implemented using XGBoost library https://xgboost.readthedocs.io/en/stable/. Training was done using 100 boosting rounds, logloss as evaluation metric, 0.1 as eta (learning rate) and setting 5 as maximum depth of each tree. The procedure was repeated 10 times.

Embeddings were obtained using Hugging Face model: https://huggingface.co/sentence-transformers/all-distilroberta-v1.

Stressor scores: We utilized a bidirectional masked language model known as Mental-RoBERTa (Ji et al., 2021), trained on a corpus collected from social forums dedicated to mental health discussions. We prompted the pretrained model with each post appending the query "Consider this post on social media to answer the question. Is the poster of this post stressed? Return Yes or No. mask ". We collected the predicted next token and its probability for inference. Additionally, we explored the experiences of the posters using a similar approach, prompting Mental-RoBERTa with: "Consider this post on social media to answer the question. The poster of this post is experiencing mask ", without guiding the model with expected responses.

Analysis using LLM: LLM analysis involved automated coding and key-phrase extraction across various dimensions. Within the coded dimensions, we inferred classifications such as determining whether the poster experienced anxiety, a panic attack, or none at all. Additionally, we classified posts into main types related to mental disorders, encompassing Personal experience, Giving information, Advocacy and awareness, Support and encouragement, Tips and strategies, as well as Question and discussion. Furthermore, we coded information related to the poster's diagnosis by a professional, ongoing treatment, and mentions of medications.

Among the other coded axes, we explored the relationship context and triggers associated with panic attacks, categorizing them into Financial stress, Argument or conflict, Traumatic experience, Social event, Life transition, Upcoming tests, Health concern, and Other. We also considered free mentions of stressors associated with these triggers.

The location (physical place) mentioned in relation to the occurrence of panic attacks was also coded into classes such as School, Work, Home, Hospital, Public spaces, Club or social venues, and Other. Additionally, we used codes for First time, Few times, and Recurrent to analyze mentions of the frequency of panic attacks.

In the realm of free-form text extraction, we focused on five axes: Mentions of physical symptoms, Cognitive symptoms, Behavioral changes, Emotional well-being, and Impact on daily life. These axes allowed us to categorize explicit or implicitly conveyed information related to the individual's experiences.

Furthermore, we extracted two additional types of shared knowledge. First, we delved into coping mechanisms individuals use to deal with panic attacks. Second, we explored perceived signs of distress that others can observe to identify when someone close is suffering from anxiety or a panic attack.

A.2 Additional material

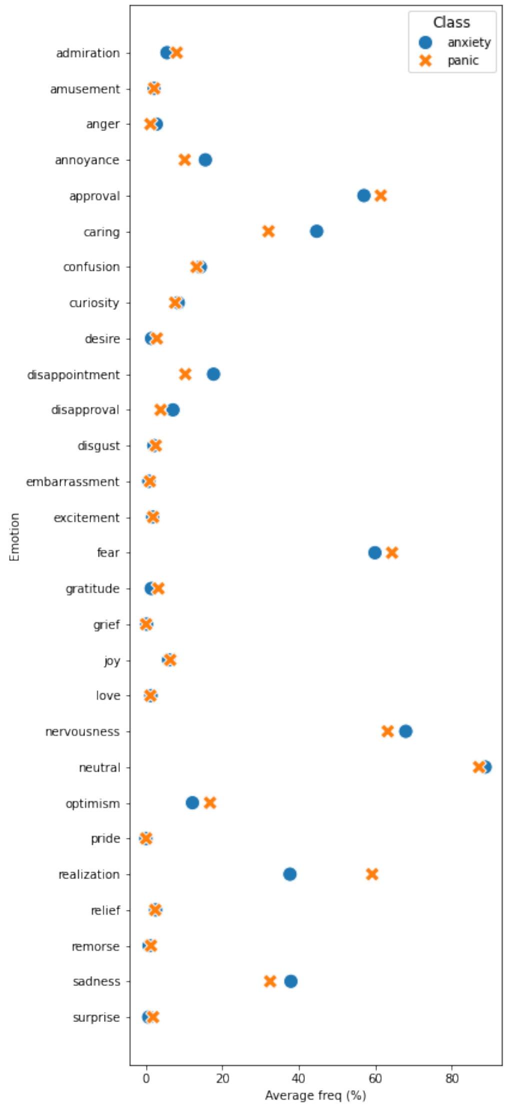

Figure 4: Average frequency of emotions, calculated based on the 5 most important emotions per post.

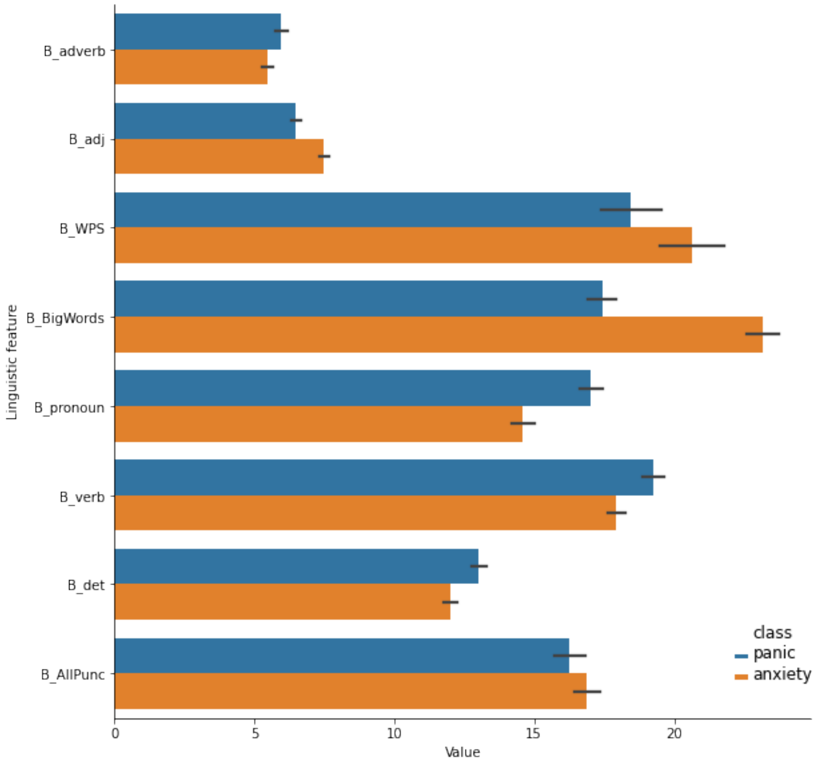

Figure 5: Linguistic features of panic and anxiety classes

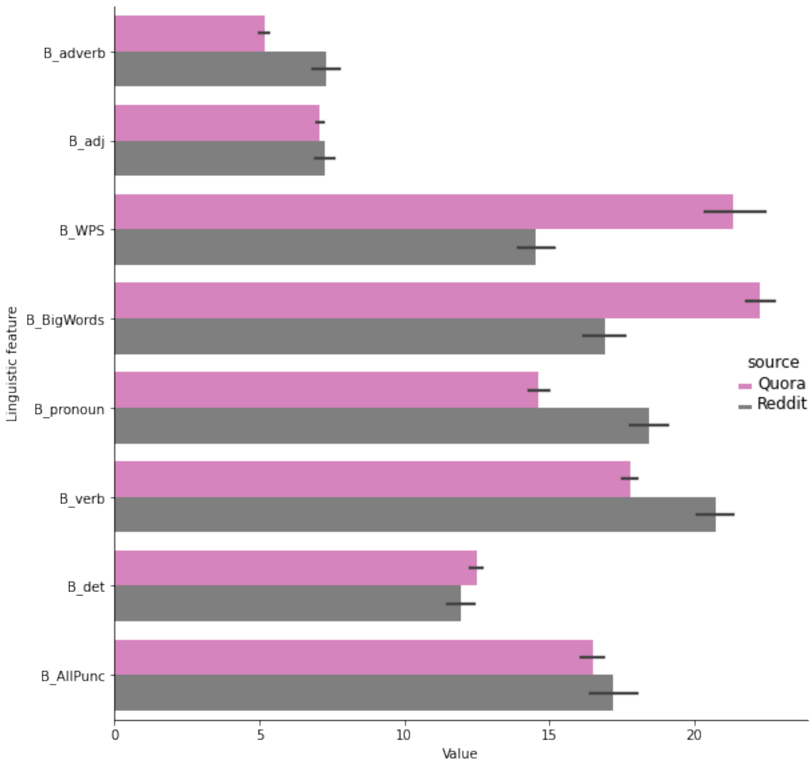

Figure 6: Linguistic features of two sources

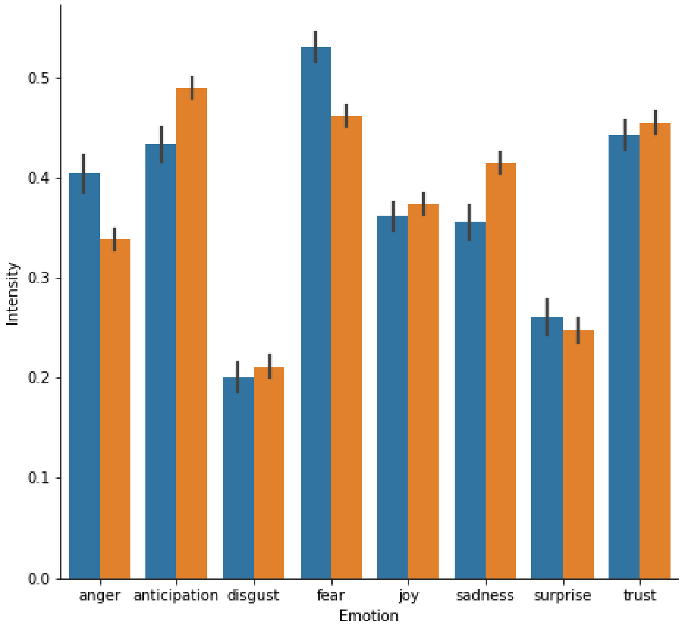

Figure 7: Emotion intensity on panic and anxiety class

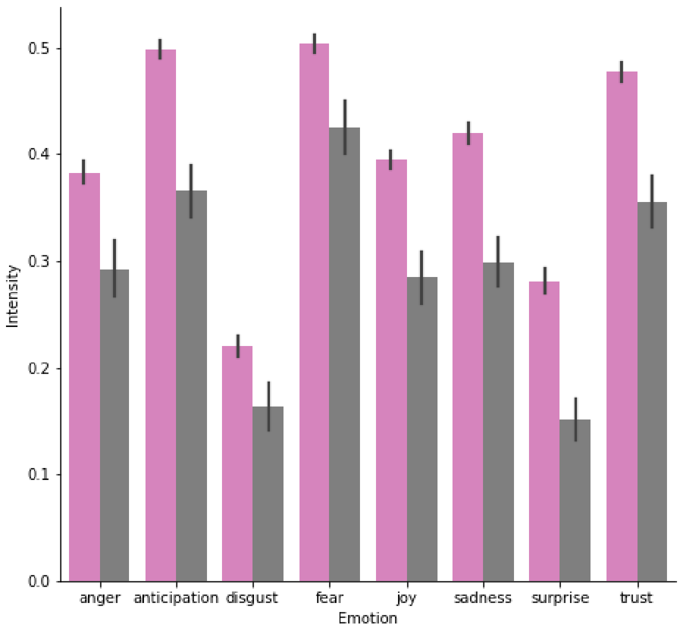

Figure 8: Emotion intensity for different sources

Category	**Question**
Panic	*What are panic attacks like?*
	Whats the worst panic attack you've ever had?
	How can I try to calm myself down while I'm having a panic attack?
	What was your first panic attack?
	What is it like to have a panic attack in public
	How does it feel like to have a panic attack at work?
	How can I tell if I had a panic attack?
	What do you do when you get panic attacks
	How to overcome panic attacks
	When was your last panic attack and how bad was it
	How-can-you-best-control-a-panic-attack-before-it-gets-out-of-control
	What is happening in the brain during a panic attack?
	Why can't I stop having panic attacks?
	How do i prevent panic attacks from happening when I'm sleeping?
	What do you take for panic attacks?
	How can I handle panic attacks at school?
	Have you ever seen someone having panic attacks?
Anxiety	*What is an anxiety?*
	Is anxiety disorder a mental illness?
	What are anxiety disorders
	Is there any difference between anxiety and social anxiety?
	Chemically and biologically what is anxiety?
	What is anxiety? What are its symptoms and its treatments?
	Why is anxiety so common today?
	What does anxiety feel like?
	What are the symptoms of anxiety
	How do you calm down when feeling anxious
	What is the real cause behind anxiety?
	How do i beat anxiety permanently?
	What did you first anxiety attack feel like?
	How do I know if I suffer from anxiety?
	How do I beat social anxiety?

Table 4: Questions used to retrieve posts per class (panic and anxiety) from two sources

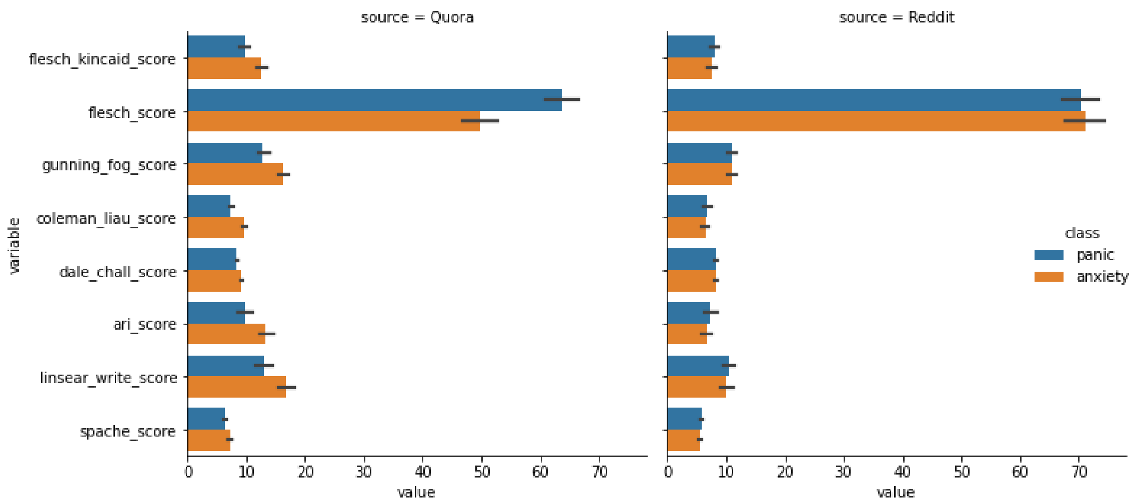

Figure 9: Different readability scores per class per source

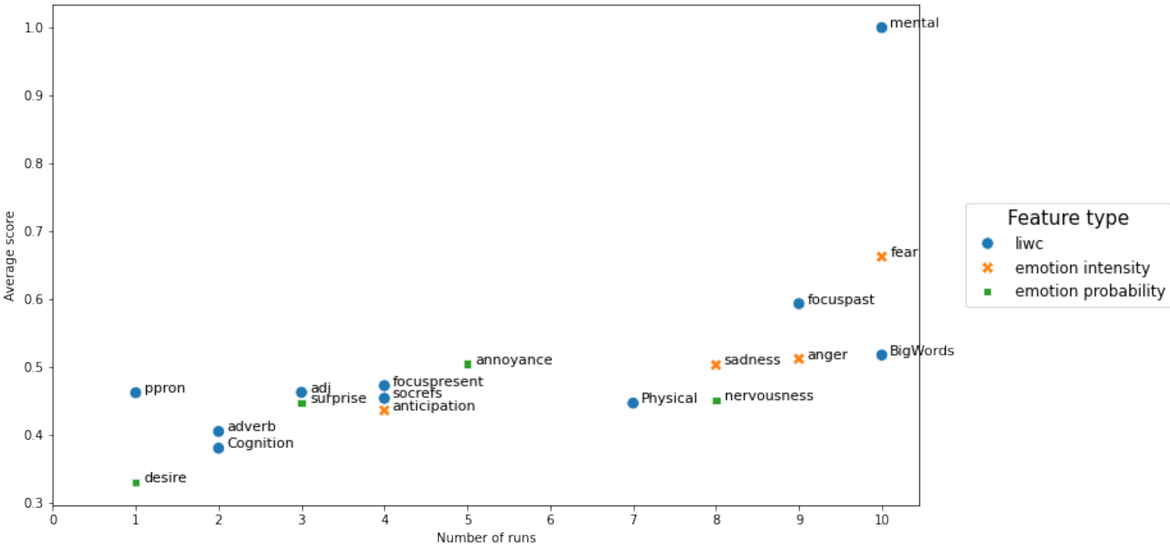

Figure 10: Most important features according to XGB feature importance across 10 runs, when considering L+M+I+W as input features. The x axis represents the number of runs where the features was considered among the top 10 most important features, while the y axis displays the average score across 10 runs.

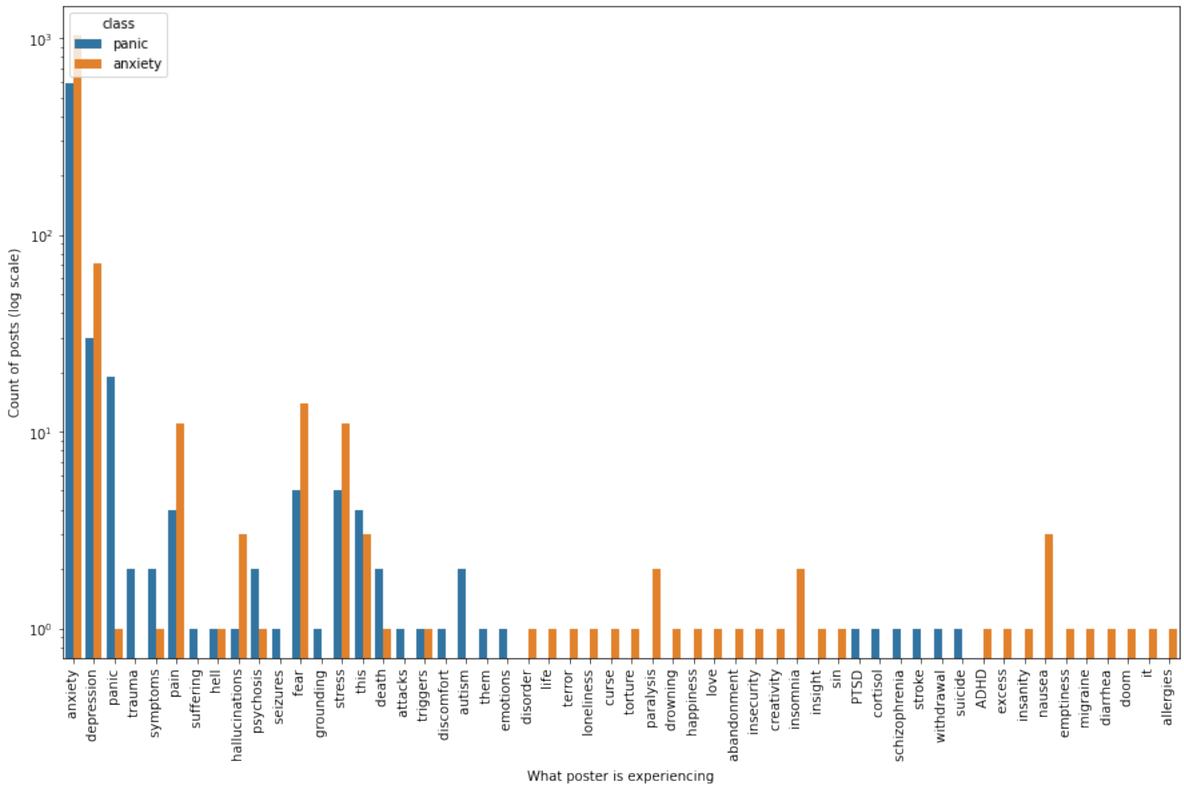

Figure 11: Variations in the distribution of anxiety and panic within the context of the poster's experiences (i.e., What is experiencing), obtained by prompting Mental_RoBERTa, are illustrated with counts presented on a logarithmic axis.

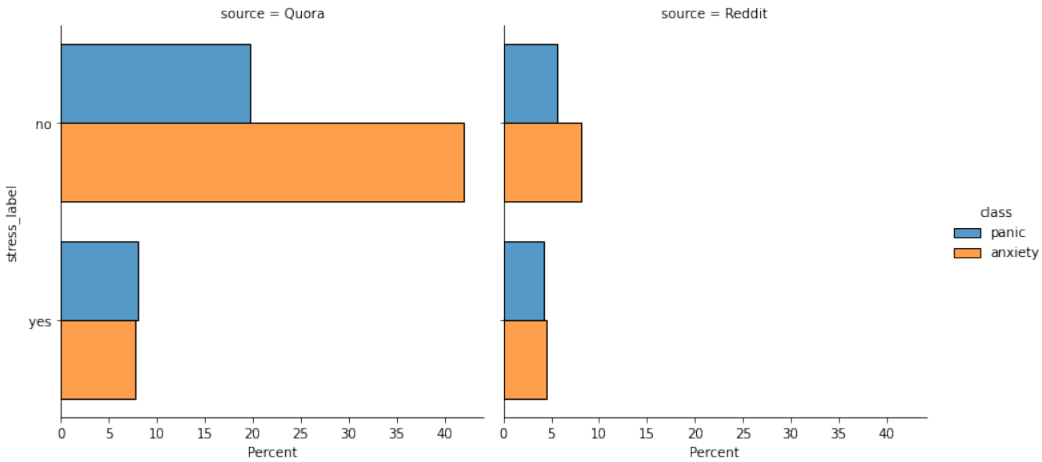

Figure 12: Stress labels per different classes and sources

Type_of_post	Counts	Response	**Anxiety** Profesionally_diagnosed	**Anxiety** Profesionally_treated	**Anxiety** Taking_medication	**Panic** Profesionally_diagnosed	**Panic** Profesionally_treated	**Panic** Taking_medication
Personal_experience	1,024	N/A	215	215	209	129	131	117
		NO	370	360	371	239	232	224
		YES	36	46	41	35	40	62
Giving_information	457	N/A	307	311	308	93	93	89
		NO	59	56	53	12	12	12
		YES	5	4	10	1	1	5
Tips_and_strategies	183	N/A	63	63	63	92	93	92
		NO	15	15	15	13	11	11
		YES	0	0	0	0	1	2
Support_and_encouragement	150	N/A	47	47	49	49	48	46
		NO	20	20	18	29	30	26
		YES	0	0	0	5	5	11
Advocacy_and_awareness	14	N/A	7	8	8	3	3	3
		NO	3	3	3	0	0	0
		YES	1	0	0	0	0	0
Question_and_discussion	17	N/A	11	11	11	1	1	1
		NO	2	2	2	3	3	3
Other	59	N/A	50	50	50	9	9	9
			Quora: 976	Reddit 241		Quora: 526	Reddit: 187	
			1,217			**713**		

Table 5: Dataset distribution with respect to post type and professional intervention, as identified by ChatGPT-3.5.