



Characterizing a New Taxonomy of Mental Disorders from Natural Language on Reddit



Scott Fleming (scottf)

Biomedical Informatics Training Program, Stanford University

WARNING: This poster contains detailed descriptions of experiences with trauma and mental illness that may be upsetting for some individuals

The Problem

- Delineations given by the *Diagnostic Statistical Manual of Mental Disorders* (DSM) yield high comorbidity, diagnostic confusion
- Shalev et al. [1998], for example, found in a prospective study that comorbid depression occurred in 44.5% of PTSD patients
- Brady et al. [2000] suggest that "high rates of comorbidity may be simply an epiphenomenon of the diagnostic criteria used"
- Grisanzio et al. [2018] characterized 6 replicable psychiatric disorder subtypes within a cohort of 420 individuals with major depressive disorder (MDD), panic disorder (PD) and posttraumatic stress disorder (PTSD)
- Does this heterogeneity extend to natural language on social media? Do the inconsistencies in subreddit labels mirror some of the known inconsistencies within the DSM-5 taxonomy? What sort of groupings would emerge from social media language itself?

The Data

- In this project, I analyze the "r/depression", "r/PanicAttack", "r/ptsd" subreddits
- These subreddits were chosen because they (1) were directly related to the MDD, PD, and PTSD diagnoses analyzed in Grisanzio et al. [2018]; (2) had large and active user bases and (3) manual review of the content suggests that users' posts are indeed related to the subreddit theme
- Using a publicly available API for querying Reddit data between January, 2006 and March, 2019, I sampled 1500 submissions from each subreddit

The Approach

- I used the second-to-last layer of a pre-trained BERT model (Devlin et al. [2018]) to obtain sentence embeddings, then averaged to obtain submission embeddings
- I also used GloVe as a comparison for featurization
- For clustering, I used the Deep Continuous Clustering (Shah and Koltun [2018]) algorithm which minimizes the following objective:

$$\mathcal{L}(\Omega, \mathbf{Z}) = \frac{1}{D} \|\mathbf{X} - G_{\omega}(\mathbf{Y})\|_F^2 + \frac{1}{d} \left(\sum_i \rho_1 (\|z_i - y_i\|; \mu_1) + \lambda \sum_{(i,j) \in \mathcal{E}} \omega_{i,j} \rho_2 (\|z_i - z_j\|; \mu_2) \right)$$

- I compared these results against K-Means Clustering as a baseline
- I used Adjusted Mutual Information (AMI) and Unsupervised Clustering Accuracy (ACC) to evaluate performance:

$$AMI(c, \hat{c}) = \frac{MI(c, \hat{c}) - E[MI(c, \hat{c})]}{\sqrt{H(c)H(\hat{c})} - E[MI(c, \hat{c})]} \quad ACC = \max_m \sum_{i=1}^n \mathbf{1}\{l_i = m(c_i)\} / n$$

Results

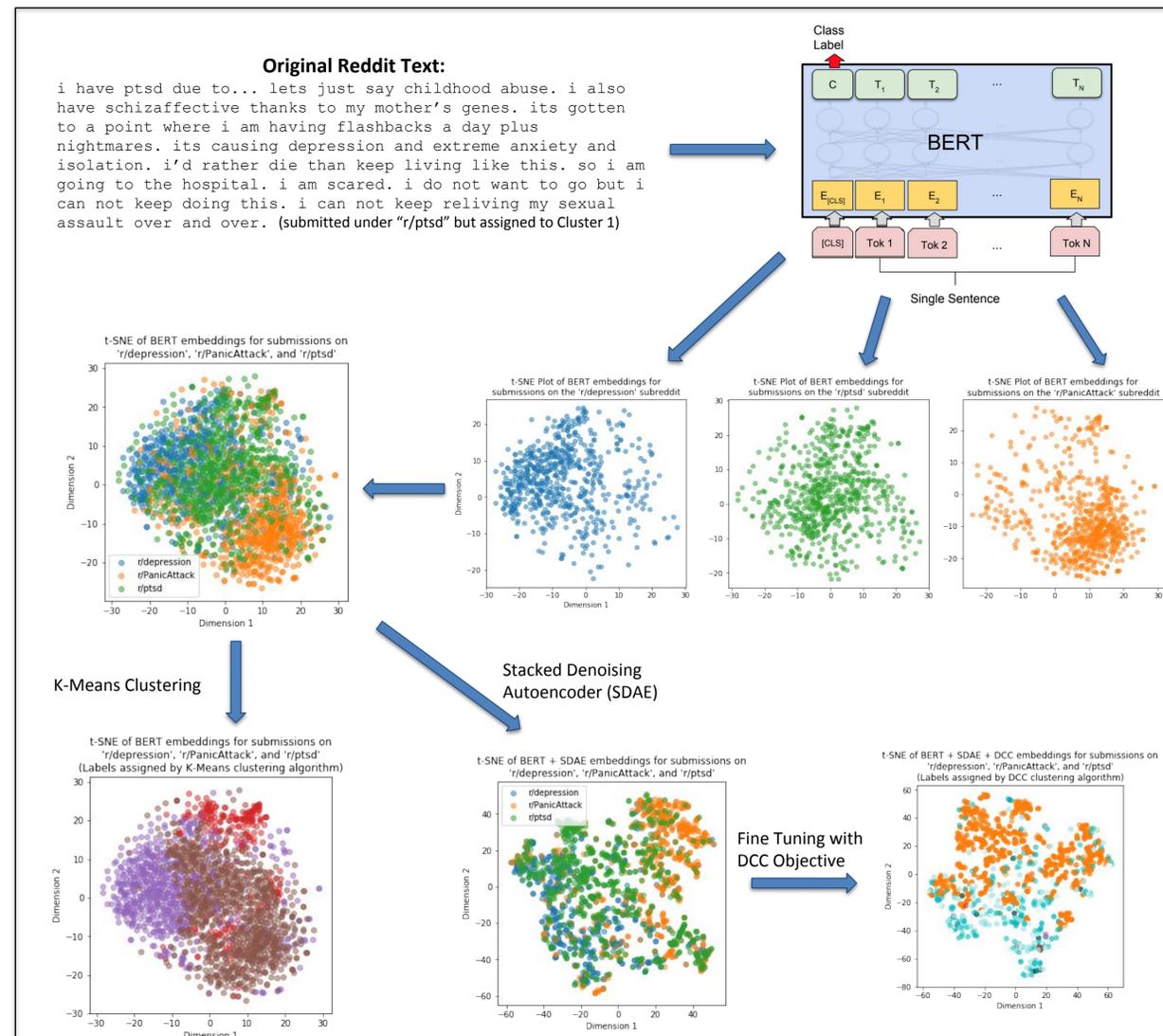
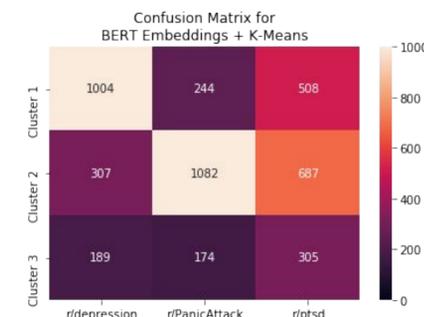
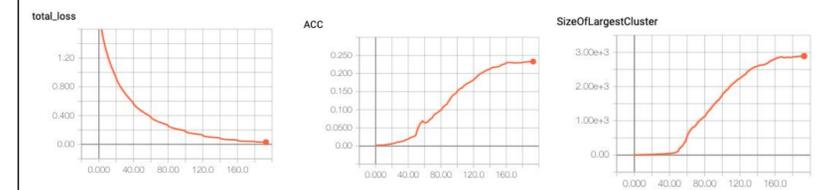


Table 1: ACC and AMI for Experiments on Reddit Data

Algorithm	AMI		ACC	
	BERT	GloVe	BERT	GloVe
K-Means	0.1023	0.0000	0.5313	0.4676
DCC	0.0383	0.0547	0.2331	0.3329



Analysis



Submitted under "r/ptsd" but assigned to Cluster 2:
so long story short i was sexually assaulted by my boyfriend at the time. i get really **bad anxiety attacks** where a symptom is where i **throw up cold sweat** or break down and cry. i was diagnosed with **major depression** 8 and trauma and i'm on lexapro and prazosin. i get nightmares or dream about him sometimes the dreams aren't very traumatic every night. i get a little **anxious** when someone mentions a certain number that is connected to the event. when i see him or get in the zone i'm not sure how to explain this i usually cannot focus and can't think straight. **is this ptsd or just trauma?**

Submitted under "r/depression" and assigned to Cluster 1:
i am a year old female. i just recently moved to vegas with my aunts and my insurance is not working so i can not get my anti depressants. i am **very suicidal**. i work around a lot of people and usually hide my **depression** very well. lately i have just sat and stared at nothing for hours. then i snap into reality and start **crying** because i realize how fucked up my emotions are. also i was abandoned by my parents and i have been wishing that i had a decent mother figure in my life. well there's my rant.

Conclusion

- Qualitative analysis suggests that mismatch of clusters to subreddit labels is indicative of a broader pattern of discordance between subreddit topic and actual submission content.
- These findings closely align with our clinical understanding of PTSD.
- These results extend our understanding of comorbidity beyond the clinical context to social media. We were not able to replicate the subtypes suggested by Grisanzio et al. [2018], however.
- The use of subreddit topics as "ground-truth" labels for predictive analyses on mental health reddit data (e.g. [Gkotsis et al., 2017]) seems inappropriate given the high rates of misalignment between subreddit label and submission content.

References

- [1] A. Y. Shalev, S. Freedman, T. Peri, D. Brandes, T. Sahar, S. P. Orr, and R. K. Pitman. Prospective study of posttraumatic stress disorder and depression following trauma. *American Journal of psychiatry*, 155(5): 630-637, 1998.
- [2] K. T. Brady, T. K. Killeen, T. Brewerton, and S. Lucerini. Comorbidity of psychiatric disorders and posttraumatic stress disorder. *The Journal of clinical psychiatry*, 2000.
- [3] K. A. Grisanzio, A. N. Goldstein-Piekarski, M. Y. Wang, A. P. R. Ahmed, Z. Samara, and L. M. Williams. Transdiagnostic symptom clusters and associations with brain, behavior, and daily function in mood, anxiety, and trauma disorders. *JAMA psychiatry*, 75(2):201-209, 2018.
- [4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805, 2018.
- [5] S. A. Shah and V. Koltun. Deep continuous clustering. arXiv preprint arXiv:1803.01449, 2018.
- [6] G. Gkotsis, A. Oelrich, S. Velupillai, M. Liakata, T. J. Hubbard, R. J. Dobson, and R. Dutta. Characterisation of mental health conditions in social media using informed deep learning. *Scientific reports*, 7:45141, 2017.

Acknowledgements

I would like to thank the NIH for its support through a BD2K Fellowship; Daniel Rubin from the Department of Biomedical Data Science; Leanne Williams from the Department of Psychiatry and Behavioral Sciences; and Imon and Nandita for suggestions.