Cluster Analysis of Online Mental Health Discourse using Topic-Infused Deep Contextualized Representations

Atharva Kulkarni[1], **Amey Hengle**[1], **Pradnya Kulkarni, Ph.D**[2], and **Manisha Marathe**[3]

[1, 3]Department of Computer Engineering, PVG's College of Engineering and Technology, affiliated to Savitribai Phule Pune University, India.
[2]Department of Psychology, Sir Parashurambhau College, Pune, India.
[1]{atharva.j.kulkarni1998, ameyhengle22}@gmail.com
[2]pradnya.kulkarni@spcollegepune.ac.in
[3]mvm_comp@pvgcoet.ac.in

Abstract

With mental health as a problem domain in NLP, the bulk of contemporary literature revolves around building better mental illness prediction models. The research focusing on the identification of discussion clusters in online mental health communities has been relatively limited. Moreover, as the underlying methodologies used in these studies mainly conform to the traditional machine learning models and statistical methods, the scope for introducing contextualized word representations for topic and theme extraction from online mental health communities remains open. Thus, in this research, we propose topic-infused deep contextualized representations, a novel data representation technique that uses autoencoders to combine deep contextual embeddings with topical information, generating robust representations for text clustering. Investigating the Reddit discourse on Post-Traumatic Stress Disorder (PTSD) and Complex Post-Traumatic Stress Disorder (C-PTSD), we elicit the thematic clusters representing the latent topics and themes discussed in the r/ptsd and r/CPTSD subreddits. Furthermore, we also present a qualitative analysis and characterization of each cluster, unraveling the prevalent discourse themes.

1 Introduction

Due to their ubiquitous nature, online health communities and social media platforms have emerged as a conducive means of information exchange and social support, especially for people with stigmatized concerns such as mental health. Consequently, these platforms provide a rich ecosystem for mental health clinicians, researchers, and practitioners to analyze the cornucopia of user-generated content and study the underlying mechanisms of different mental health conditions. With the rapid headways in artificial intelligence and computational linguistics, an increasingly large number of researchers have leveraged social media content to study various mental health illnesses and psychiatric conditions. While most of the research has focused on using the traditional machine learning models and statistical methods for predicting mental illness from social media posts, the studies addressing the discourse analysis (De Choudhury and De, 2014; Silveira Fraga et al., 2018; Loveys et al., 2018) and identification of clusters in online mental health communities (Park et al., 2018) has been relatively modest. Even with the recent surge of complex attention-based deep learning models, a large chunk of the research regarding mental health issues has focused on building better predictive systems (Benton et al., 2017; Kirinde Gamaarachchige and Inkpen, 2019; Jiang et al., 2020; Sekulic and Strube, 2019) with less emphasis on using these models for mental health related corpus analysis or information extraction.

With mental health already coalesced as an appreciable public health burden, research to investigate the discourse clusters prevalent on social media is of paramount importance. Mining information from the emergent clusters provides a lens over the dominant themes of discussion, the discourse anatomy, and the dialogue structure in the online forums while also helping to comprehend the general public engagement, sentiment, ideas, and views regarding mental health. Successful research in this direction can potentially foster identification of high-risk groups, enhanced mental health patient education programs, better diagnostic and therapeutic theory building, as well as an improved understanding of the underlying design of the online mental health communities (Park et al., 2018).

Post-traumatic stress disorder (PTSD) is a mental disorder resulting from traumatic experiences that leads to reliving the trauma, avoidance of certain situations, and hyper-vigilance. Similar to PTSD, complex post-traumatic stress disorder (C-

Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis, pages 83–93
April 19, 2021.

PTSD) is a condition that formulates the reaction resulted from the trauma, such as uncontrollable emotions, dissociation, negative self-perception, anger, mistrust, and interpersonal difficulties. Thus, in this research, we examine the online discourse on Reddit, focussing on PTSD and C-PTSD as use cases to elicit different thematic clusters present in them.

Prior research has shown that the addition of topic information to pre-trained contextualized representations yields performance improvement for semantic textual similarity (Peinelt et al., 2020). While Peinelt et al. (2020) integrated the two representations by simple concatenation, a better methodology to integrate representations from different embedding spaces is argued by Bollegala and Bao (2018). The meta-embeddings proposed by Bollegala and Bao (2018) are learned as the intermediate representations generated by various autoencoder variants. Thus, building on these two findings, we propose topic-infused deep contextualized representations, a novel data representation technique that uses a concatenated denoise autoencoder to combine deep contextual embeddings with topic information for generating robust document representations. Our methodology spawns document representations that subsume the topic information from Latent Dirichlet Allocation (Blei et al., 2003) with the contextual embeddings generated by pre-trained RoBERTa model(Liu et al., 2019). We further demonstrate that the proposed methodology achieves improvement for text clustering against the contextual embeddings generated by the pre-trained RoBERTa model.

In the light of the above discussion, our research makes the following contributions:

- We extend the methodology of word meta-embeddings to document meta-embeddings by proposing topic-infused deep contextualized representations, a data representation technique that uses a concatenated denoise autoencoder to combine deep contextual embeddings with topical information for generating robust representations for text clustering.

- We carry out a qualitative analysis and characterization of each cluster from a clinical psychology perspective.

Attribute	r/ptsd	r/CPTSD	Total
Total posts	28,133	69,600	97,733
Filtered posts	4,511	20,419	24,930
Average post length	213.52	240.36	235.51
Average comments per post	14.17	15.58	15.33
Average net upvotes per post	39.94	60.34	56.65

Table 1: Dataset statistics.

2 Related Work

Traditionally, the research in topic mining and theme extraction from online mental health communities has been focused on the use of probabilistic generative models like the Latent Dirichlet Allocation (Blei et al., 2003) and clustering techniques such as the k-means (Schütze et al., 2008). Carron-Arthur et al. (2016) employed LDA to extract topics from the internet support group BlueBoard. Results showed that users engaged in discussions with a greater topical focus on experiential knowledge, disclosure, and informational support, a pattern resembling the clinical symptom-focused approach to recovery. In their study, Dao et al. (2017) used the Hierarchical Dirichlet Process (HDP) algorithm to infer latent topics from blog posts of the LiveJournal (LJ) blogging site. The authors applied the non-parametric affinity propagation algorithm to find clusters within the online communities. Toulis and Golab (2017) compared the recurring themes encountered in private journals with the ones found in the online communities of Reddit and found significant similarities in the topics discussed across both the forums. Park et al. (2018) provide an exhaustive analysis of the thematic overlap, similarity, and difference in online mental health communities of r/depression, r/anxiety, and r/ptsd. Their results show a considerable overlap of themes between the mental health groups, attesting that people engaging in such forums face common problems and comorbidity symptoms.

Since their introduction, transformer-based language models such as BERT (Devlin et al., 2019) have led to impressive performance gains across multiple NLP tasks. Recent works show that these contextualized representations can also support type-level clusters, and hence, can be effectively used for modeling topics (Sia et al., 2020). The recent work by Thompson and Mimno (2020) demonstrates that running simple clustering algorithms

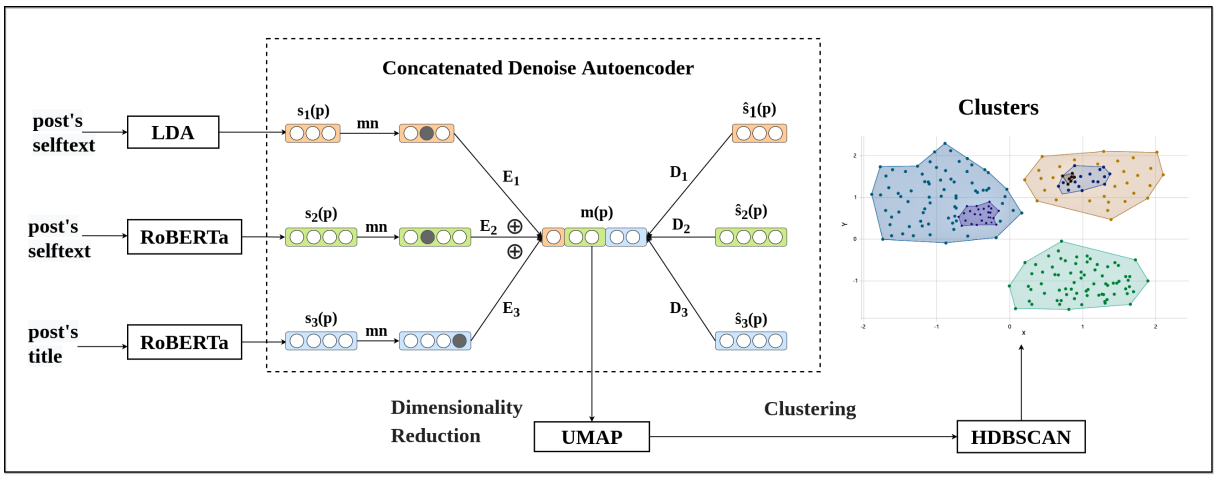

Figure 1: Architecture of the proposed system. Here, *mn* refers to masking noise applied to the autoencoder inputs.

like k-means on contextualized word representations result in word clusters that share similar properties to the ones generated by LDA. An interesting approach is followed by Peinelt et al. (2020), where the authors combine the topic models of LDA with contextualized word representations of BERT for the task of semantic similarity detection. Results depict that adding topical information improves performance, especially for examples with domain-specific terms.

3 Dataset Description

For this experiment, we selected the r/ptsd[1] and r/CPTSD [2] subreddits which have 54,000 and 97,000 active users, respectively. Using the Pushift API [3], we crawled all the available posts from these subreddits between 1st August 2015 and 31st July 2020.

To ensure that each selected post has community approval, we selected posts that have a minimum of 10 net upvotes. We further filter our dataset by eliminating posts with less than 75% English content [4], posts with less than five words, as well as posts with [DELETED], [UPDATED], and [REMOVED] entries. We employed standard text cleaning and normalization techniques for preprocessing the posts, including removing special characters, accented words, wordplays, URLs, replacing acronyms with full forms, and expanding contractions[5]. This resulted in a comprehensive dataset of 24,930 posts. The dataset statistics are provided in Table 1.

[1] https://www.reddit.com/r/ptsd/
[2] https://www.reddit.com/r/CPTSD/
[3] https://pushshift.io/
[4] https://pypi.org/project/pycld2/0.24/
[5] https://pypi.org/project/pycontractions/

4 Proposed Methodology

The proposed system consists of two key components: a robust data representation methodology and an efficient clustering algorithm. Figure 1 depicts the model architecture. Each component is elucidated in detail as follows:

4.1 Topic-infused Deep Contextualized Representations

In this section, we posit the methodology to generate the topic-infused contextualized representations, using a multi-input concatenated denoise autoencoder. The proposed autoencoder architecture has three inputs namely: the document topic distribution of the post's selftext [6], contextualized document embedding of the post's selftext, and the contextualized document embedding of the post's title. Let S_1, S_2, and S_3 denote the corresponding three input embedding spaces of dimensions d_1, d_2, and d_3, respectively. Let N be the total number of posts. For each post $p \in N$, the three document embeddings are given by $s_1(p) \in \mathbb{R}^{d_1}$, $s_2(p) \in \mathbb{R}^{d_2}$, and $s_3(p) \in \mathbb{R}^{d_3}$. The autoencoder model consists of three encoders E_1, E_2, and E_3 which encode the source embeddings to a common meta-embedding space M of dimensionality d_m. Each encoder independently performs dimensionality reduction and non-linear transformations on the respective embeddings, thus, learning to retain essential information from each source embedding. Dimensionalities of the encoded input embeddings are denoted respectively, by d'_1, d'_2,

[6] selftext attribute refers to the main body of the post.

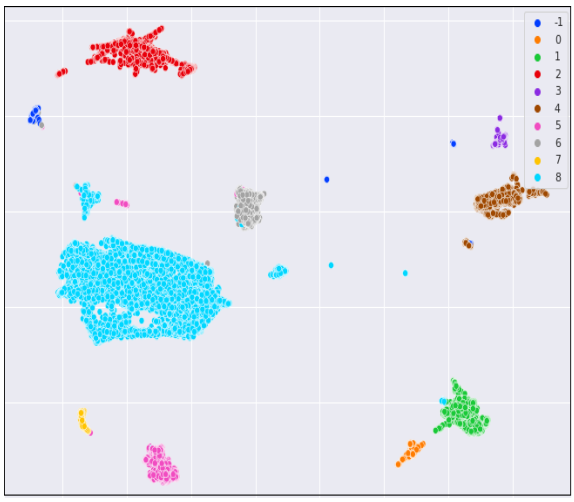

Figure 2: 2D embdding space of the topic-infused deep contextualized representations (-1 represents the unassigned datapoints).

and d'_3. The concatenation of each of the encoded input source embeddings results in the document meta-embedding $m(p)$, as given by Equation 1.

$$m(p) = E_1(s_1(p)) \oplus E_2(s_2(p)) \oplus E_3(s_3(p)) \quad (1)$$

Therefore, the dimensionality of document meta-embedding space M is computed as the sum of dimensionalities of encoded source embeddings. It is given by Equation 2.

$$d_m = d'_1 + d'_2 + d'_3 \quad (2)$$

The three decoders D_1, D_2, and D_3, try to reconstruct the individual source embeddings from the document meta-embeddings, thereby implicitly utilizing the common and the complementary information present in the source embeddings. Equations 3, 4, and 5 represent the reconstructed versions of the source embeddings, given by $\hat{s}_1(p)$, $\hat{s}_2(p)$, and $\hat{s}_3(p)$.

$$\hat{s}_1(p) = D_1(m(p)) \quad (3)$$

$$\hat{s}_2(p) = D_2(m(p)) \quad (4)$$

$$\hat{s}_3(p) = D_3(m(p)) \quad (5)$$

The objective loss $\mathcal{L}$ is calculated as the sum of the reconstruction loss for each of the three input embeddings. It is formulated in Equation 6.

$$\mathcal{L} = \sum_{p \in N} (||\hat{s}_1(p) - s_1(p)||^2 + ||\hat{s}_2(p) - s_2(p)||^2 + ||\hat{s}_3(p) - s_3(p)||^2) \quad (6)$$

Cluster	Original cluster size	Exemplar cluster size
Cluster 1	360	300
Cluster 2	1756	300
Cluster 3	2829	600
Cluster 4	361	300
Cluster 5	1971	300
Cluster 6	1495	300
Cluster 7	1348	300
Cluster 8	417	300
Cluster 9	14065	2100
Unassigned	46	-

Table 2: Cluster and respective exemplar sizes.

Thus, the proposed system jointly learns E_1, E_2, E_3, and D_1, D_2, D_3, such that the loss given by Equation 6 is minimized.

4.2 Dimensionality Reduction and Clustering

In this study, we make use of HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise), a hierarchical density-based unsupervised clustering technique to congregate semantically-similar posts together in clusters (McInnes et al., 2017). Unlike partition-based clustering algorithms, HDBSCAN can find varying density clusters and is more robust to parameter selection. Moreover, HDBSCAN does not force the data points to belong to any particular cluster, making it suitable for handling outliers and noisy data.

As HDBSCAN uses relative-distance measures for clustering, it often suffers from the curse of dimensionality (McInnes et al., 2017). As the dimension of the topic-infused contextualized embeddings is quite high (d_m = 768), we employ UMAP (Uniform Manifold Approximation Projection), a technique for general non-linear dimensionality reduction (McInnes et al., 2018). UMAP is preferred over other dimensionality reduction algorithms as it keeps a significant portion of the high-dimensional local structure in lower dimensionality, thus, causing minimal information loss.

5 Experimental Setup

The document topic distributions are generated using the LDA mallet python version [7]. As the topics of interest centers around entities that are mostly

[7] https://radimrehurek.com/gensim/models/wrappers/ldamallet.html

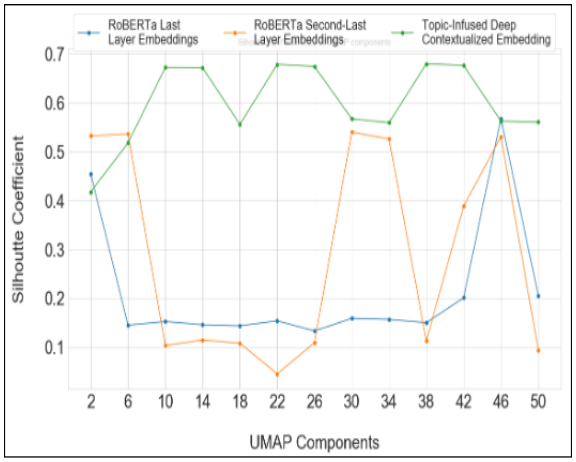

Figure 3: Silhoutte Coefficient (SC) Comparison

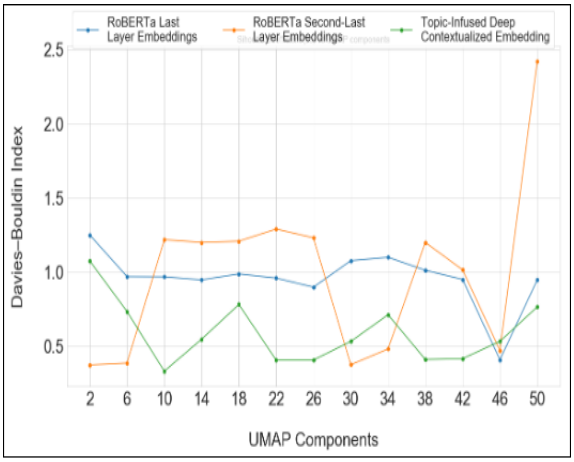

Figure 5: Davies Bouldin Index (DBI) Comparison

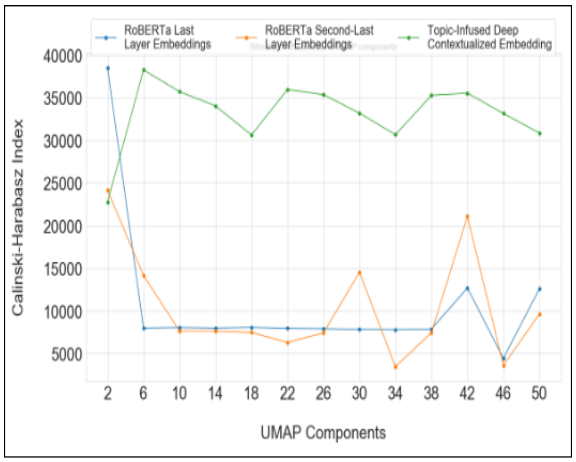

Figure 4: Calinski Harabasz Index (CHI) Comparison

nouns, we follow a nouns-only approach as employed by Martin and Johnson (2015) for topic modeling. The number of topics is empirically chosen as 10 since it displayed the best topic coherence score (c_v) of 0.49. The rest of the hyperparameters for LDA are kept at default. The 768-dimensional contextualized document embeddings are generated as the average of all the embeddings for each word in the document, extracted from the second last layer of the pre-trained RoBERTa-base model (Liu et al., 2019). Thus, the three input d_1, d_2, and d_3 are of dimensions 10, 768, and 768, respectively.

In our experiments, each autoencoder is implemented as a single hidden layer neural network. The dimensions d_1', d_2', and d_3' are chosen as 10, 379, and 379, respectively. The hidden dimensions are chosen as such so that the topic-infused deep contextualized representations are of dimensions 768, making them comparable with that of RoBERTa. Masking noise of 10 percent is applied to the source embeddings before encoding (Vincent et al., 2010). Leaky rectified linear (Leaky ReLU) (Maas et al., 2013) activation is applied to each layer with the default parameter setting. The model is trained end-to-end for 200 epochs, with the Adam optimizer (Kingma and Ba, 2015), a learning rate of 0.001, and a mini-batch size of 128 for minimizing the objective loss. The learning rate is reduced by a factor of 0.1 if validation loss does not decline after 10 successive epochs. The model with the best loss is used for prediction.

6 Clustering Performance Evaluation

In order to assess the clustering performance, we make use of three measures, namely the Silhouette Coefficient (SC) (Rousseeuw, 1987), the Calinski-Harabasz Index (CHI) (Caliński and Harabasz, 1974), and the Davies-Bouldin Index (DBI) (Davies and Bouldin, 1979). Higher SC, and CHI, and lower DBI scores indicate a better separation of the clusters and tighter integration inside the clusters. We compare the clustering performance of three embedding sets, namely: the RoBERTa last layer embeddings, the RoBERTa second last layer embeddings, and our proposed topic-infused deep contextualized embeddings. Figures 3, 4, and 5 depict the performance comparison of the corresponding three embeddings with respect to the three metrics mentioned above. The y-axis represents the three corresponding metrics, whereas the x-axis represents the UMAP component variations. The hyperparameter values of the number of neighbours and minimum distance were chosen as 30 and 0.0, respectively, while cosine similarity was used as the metric for computing the distance in the ambient space of the input data. The random state is seeded to an integer value (42). For HDB-

SCAN, the minimum cluster size was selected as 300, with the other parameters set to the default values (McInnes et al., 2017). From the comparative analysis, it is evident that the proposed topic-infused deep contextualized representations result in an improved clustering performance across all three metrics. For our study, we choose UMAP with components 10, as it empirically gives consistent performance across all three metrics. The 2D embedding space after clustering is shown in Figure 2.

To reduce the cluster size, we extract the most representative posts or exemplars of each cluster [8]. In technical terms, exemplars are the data points that lie at the heart of a cluster, around which the ultimate cluster forms. Table 2 provides the statistics of the clusters and their respective exemplar sizes. To find the key-words for each cluster, we employ a class-based term-frequency inverse-document-frequency (c-TF-IDF)[9] method on each of the obtained exemplars. Unlike the traditional TF-IDF, which considers each document of a corpus, c-TF-IDF is a class-based method that treats all the documents belonging to a particular class as a single document. This enables us to find only the latent topics most representative of a particular cluster and penalize the frequent words across the clusters. For each cluster, the c-TF-IDF score is calculated using the Equation 7, where each word t is extracted for each class i, and the number of documents m is divided by the total frequency of the word t across all classes n.

$$c - TF - IDF_i = \frac{t_i}{w_i} * \log \frac{m}{\sum_{j=1}^{n} t_j} \quad (7)$$

The coherence scores for the topics generated by LDA and the clusters generated by the above-mentioned embeddings are reported in Table 3. The top 15 nouns with the highest c-TF-IDF scores for each cluster are used to evaluate the topic coherence. The overall coherence of the topics generated is evaluated using extrinsic as well as intrinsic measures. As the intrinsic topic coherence is computed using word co-occurrences in documents from the corpus (Mimno et al., 2011), it is natural that LDA reports the highest intrinsic coherence (c_v), as it is based on word-occurrences statistics. Moreover, as argued by Stevens et al. (2012), a good intrinsic coherence score does not necessarily guarantee that the generated topics make semantic sense or that they are interpretable by humans. *Normalized pointwise mutual information* (NPMI) (Bouma, 2009), on the other hand, has been shown to correlated with human judgement (Lau et al., 2014). Thus, in this study, we use NPMI as an extrinsic measure. Experiments by Sia et al. (2020) suggest that clustering contextual embeddings can result in topics with better NPMI compared to LDA. As evident from the results reported in Table 3, our results corroborate these findings, as our proposed methodology exhibits the best NPMI score. Figure 6 depicts the wordclouds of the top 50 nouns with the highest c-TF-IDF scores for each cluster.

Embeddings	**Extrinsic Measure**	**Intrinsic Measure**
RoBERTa last layer	-0.0678	0.2650
RoBERTa 2nd last layer	-0.0500	0.2702
Topic-infused deep contextualized representations	**-0.0255**	0.3186
LDA	-0.2994	**0.4972**

Table 3: Topic coherence results.

7 Thematic Analysis of the Clusters

For a comprehensive qualitative evaluation, we select the top 100 posts from each clusters which exhibit the highest Jaccard Index scores [10] with that of the top 15 nouns with the highest c-TF-IDF scores for each cluster. The qualitative analysis is carried out by the subject expert in clinical psychology to draw out the major discussion themes. Total 7 themes were extracted from the 9 clusters generated. They are as follows:

Abuse stories: Cluster 3 reveals self-disclosure stories of abuse that are physical, emotional, or sexual in nature. Abusers are mainly parents, siblings, close family members, co-workers, and known people. Narrations prominently highlight dysfunctional family dynamics during the victim's childhood as well as the issues related to current family interactions. Family dynamics are characterized by alcoholic/ substance abusive parents and abusers with pathological personality traits. Posts indicate traumatic experiences, the effect of these experiences, and how they act as triggers. The victim's exposure to the trauma was long and frequent.

[8] https://hdbscan.readthedocs.io/en/latest/soft_clustering_explanation.html

[9] https://github.com/MaartenGr/cTFIDF

[10] https://en.wikipedia.org/wiki/Jaccard_index

Figure 6: Wordclouds of top 50 c-TF-IDF nouns for the 9 clusters.

> *"My **mother** was very **controlling, scary, abusive and manipulative**..."*
>
> *"I've been coming to terms over the past 6 months or so with the fact that my **family** caused me a significant amount of **trauma growing up**..."*
>
> *"For as long as I can remember, my **father** was **physically and emotionally abusive**..."*

Flashbacks: The primary themes of cluster 6 include flashbacks of traumatic incidences and the emotional disturbances associated with them. Feelings of anxiety, panic, nightmares, fear, impulsivity, and anger are prevalent throughout the cluster. Narrations reveal the way flashbacks are impeding current life issues. Examples include an aversion to touch and sexual experiences, difficulties in romantic relationships, social relationships, daily chores, and work as well as triggering health-related symptoms. Working through flashbacks, ways of dealing with it, and help to overcome it are also shared.

> *"...struggling with **flashbacks, nightmares, and intrusive thoughts** about **trauma**. can anyone give me reassurance that i'll get through this?..."*

> *"...I can stay that way for weeks while I slowly process the **old trauma** that comes up. I'm in that state right now - just **overwhelmed** by **intense emotional flashbacks**..."*

Advice seeking: People have vividly expressed their feelings and are seeking advice about various issues related to PTSD, according to the posts in cluster 8. Main themes revolve around advice related to job functioning impacted by PTSD symptoms, social isolation, dependency issues, management of difficult emotions, exhaustion, and frustration. Posts in this cluster are brief and direct, seeking advice, help, and support from the Reddit community.

> *"...I would appreciate any **tips or advice**. **How do you** deal with emotional isolation?"*
>
> *"When I'm more myself, I feel like I'm too cynical and rude. **Anyone else** have experience with this? **How do you** become more authentic? "*

Therapy and therapist experiences: Cluster 1 and 4 mainly focuses on therapy and therapist experiences. Posts highlight the process of therapy, experiences with therapists, challenges faced in

the therapeutic process, insights from therapies, transference experiences, and seeking help for such issues. Other findings include posts that seek therapeutic methods other than visiting therapists. Surprisingly, a large chunk of the posts portrays a negative connotation about therapy experiences or therapy.

> *"I've been **traumatized by therapists** in my childhood that relayed info back to my abusive parents and gave them more ammo to hurt me with..."*
>
> *"The only **therapist** who is willing to treat me is **unreliable, too far away, and frankly unsympathetic and unqualified**..."*

Difficulty in emotional regulation: Posts in cluster 5 primarily focus on different emotions associated with trauma. Posts suggest difficulty in controlling emotions, emotional neglect, panic, anxiety, feeling of emptiness, hopelessness, emotional distancing due to trauma, difficulty understanding and expressing emotions, and ambivalent emotions towards abusers. Other findings include the progress in trauma-related emotions. Most of the posts are help-seeking in nature, for validating their emotions.

> *"DAE **struggle** with identifying or verbally explaining your **emotions**?"*
>
> *"I have recently realized that I have the **hardest time identifying my emotions**..."*
>
> *"DAE struggle with **guilt** around cutting off **toxic people**?"*
>
> *"How did you overcome chronic **emotional numbness**? Inability and lack of desire to develop a connection with others (platonic or romantic)?"*
>
> *"DAE feel like they've **built a wall** between themselves and their **feelings**?"*

Working through traumatic abuse: Post in cluster 2 narrates retrospective abusive, traumatic experiences and their psychological aftereffects. The majority of the posts indicate signs of shame, dissociation, social withdrawal, anxiety, self-downing, and victimization. Confrontation, triumph while struggling with vulnerabilities, hope, and new insights about oneself are other findings.

> *"Mainly the reasons I **self-isolate, distrust/avoid** intimacy, and deprive myself of certain pleasures is due to my PTSD symptoms because of all the **abuse** I endured..."*
>
> *"The lifelong **shame and fear**, all the problems causing me to not succeed in life or have any healthy relationships..."*
>
> *"I haven't lived a day in my adult life. It's all been a decades-long **dissociation**."*

Abuse issues and general PTSD: Clusters 7 and 9 mostly contains posts about abuse stories and PTSD in general. The major themes are related to the parent-child issues, insecure parenting styles, emotional and physical abuse, the effect of childhood trauma, loss and grieving for not having healthy parental relations, the impact of insecure parenting, disclosure of trauma incidences, and sharing recovery tips. Following are a few illustrations:

> *".. I **hate** the days where you long for the **childhood** that you never had and **support you never received**. For the belonging and "home" that you never had ..."*
>
> *"... My last therapy session had me divulging my **experiences as a young child** and what life was like for me back then..."*

8 Discussions

Cluster structure is in accordance with the formal diagnostic criteria of PTSD and C-PTSD. Abuse stories reflected in cluster 1 corresponds to the *criterion-A* of PTSD in DSM-5 (Association et al., 2013), that states the experience of trauma. Out of the six symptom cluster proffered by ICD-11[11] for CPTSD (re-experiencing, avoidance, hypervigilance, emotional dysregulation, interpersonal difficulties, and negative self-concept), cluster 6 adequately corresponds to re-experiencing of flashbacks, while cluster 5 is consistent with emotional dysregulation. Parent-child relationships portrayed in cluster 2 corresponds to etiological factors. Dysfunctional parent-child relation in PTSD has been widely confirmed in the clinical psychology literature (Cockram et al., 2010; Cross et al., 2018; van Ee et al., 2016). On the other hand, clusters 5, 6, and 7 characterizing therapy experience, working

[11] https://icd.who.int/en

through trauma, and advice, respectively, attribute to the treatments and interventions aspect of PTSD. In conclusion, it can be said that these clusters reveal salient features of PTSD.

Although the clusters more or less convey their discrete independent themes, there exist many common topics running throughout the clusters. The recurrent expressions observed throughout the clusters pertain to **personal and social life** *(family, parents, friends, person, relationships)*, **daily grind** *(work, home, job, place)*, and **temporal indicators** *(day, time, year, month, today, week)*. Furthermore, words encompassing **cognition** *(thoughts, think, feel, feeling, lot, hard)*, **emotional and affect expressions** *(happy, love, good, anxiety, kind, bad, hate)*, and **inhibition expressions** *(avoid, deny, escape)* are perpetual throughout the posts.

9 Limitations

Our work has its own drawbacks and limitations. The problem statement at hand is of fine-grained clustering rather than coarse-grained clustering, thus, making it difficult to draw out interpretable topics of discussion. This is attested by the low NPMI scores in Table 3. Furthermore, the posts' lengthy nature (refer Table 1) makes it difficult for models like RoBERTa to capture maximum contextual information (Liu et al., 2019). Many psychiatric disorders are found to be co-morbid with PTSD (Brady et al., 2000). The research by Park and Conway (2018) indicates that people facing mental health issues often find it difficult to regulate their ideas and views. Their research further entails that posts from online mental health communities are more difficult to read and portray less lexical diversity. These factors further hinder the elicitation of interpretable topics from the corpus. Also, not all clusters exhibit independent themes of discussions and some themes are a combination of multiple clusters.

10 Conclusion and Future Work

In this research, we introduce topic-infused deep contextualized representations, a robust data representation methodology that successfully integrates topical information with contextualized embeddings. We collect and analyze large-scale data pertaining to the online discourse on Reddit, centered around PTSD and CPTSD, and perform density-based clustering to draw out the prominent clusters and analyze the themes of discussion present in them. Despite the perpetual semantic and thematic similarities amongst the posts in the corpus, our methodology, to some extent, is able to draw out the underlying, fine-grained, latent clusters. The qualitative analysis of each cluster revealed the characteristic themes and salient features of PTSD and CPTSD, consistent with the clinical psychology literature. Through the lens of social media, this study delineates a deeper understanding of PTSD and C-PTSD, fostering further research in early detection of mental illnesses, identification of high-risk groups, enhanced mental health patient education programs, better diagnostic and therapeutic theory building, as well as an improved understanding of the underlying design of the online mental health communities.

The future work of this research can take multiple directions. From an NLP standpoint, the robustness of the proposed methodology should be examined by testing them on various other benchmark NLP tasks such as semantic textual similarity, word analogy, and text classification, to name a few. Other variants of autoencoders and objective losses could be employed to facilitate tighter integration of topical information with the contextualized embeddings. From the mental health and clinical psychology perspective, such research can be easily extended to other online mental health communities to draw useful insights.

References

American Psychiatric Association et al. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Adrian Benton, Margaret Mitchell, and Dirk Hovy. 2017. Multitask learning for mental health conditions with limited social media data. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 152–162, Valencia, Spain. Association for Computational Linguistics.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Danushka Bollegala and Cong Bao. 2018. Learning word meta-embeddings by autoencoding. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1650–1661, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Gerlof Bouma. 2009. Normalized (pointwise) mutual

information in collocation extraction. *Proceedings of GSCL*, pages 31–40.

Kathleen T Brady, Therese K Killeen, Tim Brewerton, and Sylvia Lucerini. 2000. Comorbidity of psychiatric disorders and posttraumatic stress disorder. *The Journal of clinical psychiatry*, 61(suppl 7):22–32.

T. Caliński and J Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics*, 3(1):1–27.

Bradley Carron-Arthur, Julia Reynolds, Kylie Bennett, Anthony Bennett, and Kathleen M Griffiths. 2016. What's all the talk about? topic modelling in a mental health internet support group. *BMC psychiatry*, 16(1):367.

David M Cockram, Peter D Drummond, and Christopher W Lee. 2010. Role and treatment of early maladaptive schemas in vietnam veterans with ptsd. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice*, 17(3):165–182.

Dorthie Cross, L Alexander Vance, Ye Ji Kim, Andrew L Ruchard, Nathan Fox, Tanja Jovanovic, and Bekh Bradley. 2018. Trauma exposure, ptsd, and parenting in a community sample of low-income, predominantly african american mothers and children. *Psychological trauma: theory, research, practice, and policy*, 10(3):327.

Bo Dao, Thin Nguyen, Svetha Venkatesh, and Dinh Phung. 2017. Latent sentiment topic modelling and nonparametric discovery of online mental health-related communities. *International Journal of Data Science and Analytics*, 4(3):209–231.

D. L. Davies and D. W. Bouldin. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227.

Munmun De Choudhury and Sushovan De. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1).

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Elisa van Ee, Rolf J Kleber, Marian J Jongmans, Trudy TM Mooren, and Dorothee Out. 2016. Parental ptsd, adverse parenting and child attachment in a refugee sample. *Attachment & human development*, 18(3):273–291.

Zhengping Jiang, Sarah Ita Levitan, Jonathan Zomick, and Julia Hirschberg. 2020. Detection of mental health from Reddit via deep contextualized representations. In *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis*, pages 147–156, Online. Association for Computational Linguistics.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Prasadith Kirinde Gamaarachchige and Diana Inkpen. 2019. Multi-task, multi-channel, multi-input learning for mental illness detection using social media text. In *Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019)*, pages 54–64, Hong Kong. Association for Computational Linguistics.

Jey Han Lau, David Newman, and Timothy Baldwin. 2014. Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 530–539, Gothenburg, Sweden. Association for Computational Linguistics.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Kate Loveys, Jonathan Torrez, Alex Fine, Glen Moriarty, and Glen Coppersmith. 2018. Cross-cultural differences in language markers of depression online. In *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*, pages 78–87, New Orleans, LA. Association for Computational Linguistics.

Andrew L. Maas, Awni Y. Hannun, and Andrew Y. Ng. 2013. Rectifier nonlinearities improve neural network acoustic models. In *in ICML Workshop on Deep Learning for Audio, Speech and Language Processing*.

Fiona Martin and Mark Johnson. 2015. More efficient topic modelling through a noun only approach. In *Proceedings of the Australasian Language Technology Association Workshop 2015*, pages 111–115, Parramatta, Australia.

Leland McInnes, John Healy, and Steve Astels. 2017. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11).

Leland McInnes, John Healy, and James Melville. 2018. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

David Mimno, Hanna Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 262–272, Edinburgh, Scotland, UK. Association for Computational Linguistics.

Albert Park and Mike Conway. 2018. Harnessing reddit to understand the written-communication challenges experienced by individuals with mental health disorders: Analysis of texts from mental health communities. *J Med Internet Res*, 20(4):e121.

Albert Park, Mike Conway, and Annie T. Chen. 2018. Examining thematic similarity, difference, and membership in three online mental health communities from reddit. *Comput. Hum. Behav.*, 78(C):98–112.

Nicole Peinelt, Dong Nguyen, and Maria Liakata. 2020. tBERT: Topic models and BERT joining forces for semantic similarity detection. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7047–7055, Online. Association for Computational Linguistics.

Peter Rousseeuw. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20(1):53–65.

Hinrich Schütze, Christopher D Manning, and Prabhakar Raghavan. 2008. *Introduction to information retrieval*, volume 39. Cambridge University Press Cambridge.

Ivan Sekulic and Michael Strube. 2019. Adapting deep learning methods for mental health prediction on social media. In *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pages 322–327, Hong Kong, China. Association for Computational Linguistics.

Suzanna Sia, Ayush Dalmia, and Sabrina J. Mielke. 2020. Tired of topic models? clusters of pretrained word embeddings make for fast and good topics too! In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1728–1736, Online. Association for Computational Linguistics.

B. Silveira Fraga, A. P. Couto da Silva, and F. Murai. 2018. Online social networks in health care: A study of mental disorders on reddit. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 568–573.

Keith Stevens, Philip Kegelmeyer, David Andrzejewski, and David Buttler. 2012. Exploring topic coherence over many models and many topics. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 952–961, Jeju Island, Korea. Association for Computational Linguistics.

Laure Thompson and David Mimno. 2020. Topic modeling with contextualized word representation clusters. *arXiv preprint arXiv:2010.12626*.

Andrew Toulis and Lukasz Golab. 2017. Social media mining to understand public mental health. In *VLDB Workshop on Data Management and Analytics for Medicine and Healthcare*, pages 55–70. Springer.

Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, 11(110):3371–3408.