Topic Modeling on Mental Health-Related Tweets

1. Introduction

The digital era uses Twitter as a forceful resource that enables users to share their life stories and express viewpoints about various subjects, including mental wellness issues. These digital platforms make instant communication possible because they give users an open environment to share their mental health challenges, receive support, and educate others about psychological health conditions [1]. The research utilizes Natural Language Processing (NLP) and Machine Learning (ML) to evaluate mental health discussions on Twitter. Through Latent Dirichlet Allocation (LDA) and BERTopic, the study investigates predominant subject matters by monitoring public sentiment and tracking recurring discussion patterns [2]. The investigation process begins with data collection and preprocessing, followed by topic modeling techniques, and concludes with visualization methods to effectively analyze mental health discussions [3].

2. Data Collection and Preprocessing

2.1 Dataset Selection

The research dataset consists of tweets about mental health obtained from Kaggle[4]. The selection of this dataset with 20,000 tweets was made to achieve a complete and unbiased analysis. The metadata fields include:

- **Tweet ID:** It provides individual identification tags which make tweet tracking and reference possible.
- **Timestamp:** It includes the precise time stamp when every posting occurred thus enabling scientists to analyze temporal data patterns.
- **Post Text:** The basis of the topic modeling approach lies with tweet post text which represents the original written content of the posts.
- **User Information:** The analysis evaluates tweet influence through information about users which includes their IDs in addition to follower numbers and engagement metrics.
- **Engagement Metrics**: Interaction statistics, including retweet counts and likes, providing insights into the popularity of discussions.

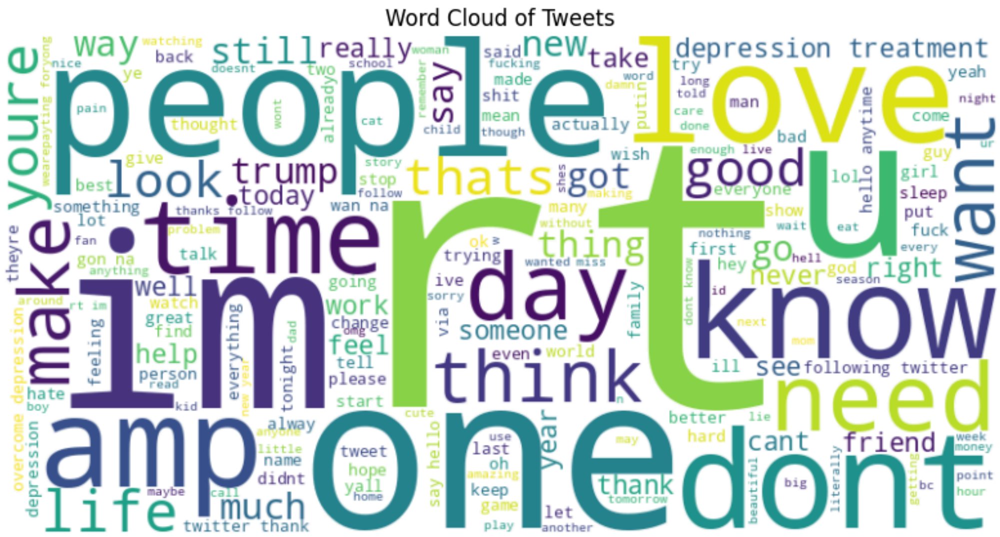

Figure 1: Word Cloud of the Dataset

2.2 Data Preprocessing Steps

2.2.1 Text Cleaning:

A complete text-cleaning pipeline standardized the dataset by implementing these steps-

- **Lowercasing:** Everything was converted to lowercase to keep things uniform and avoid duplicates.
- **Removing URLs & Mentions:** Links and @usernames were deleted since they don’t add meaning.
- **Hashtag & Special Character Removal:** Hashtags, punctuation, and symbols were stripped to focus on the core text.
- **Cleaning Up Language & Whitespace:** Non-English words were removed for consistency, and extra spaces were cleaned up.
- **Filtering Retweets:** Retweets were excluded to keep only original content and avoid duplication.

2.2.2 Tokenization and Stopword Removal:

After the text was cleaned, it was broken down into tokens (words). This was achieved using NLTK’s word tokenizer, which effectively segments sentences into discrete words. To further refine the text:

- **Common stopwords** (e.g., "the," "and," "is") were removed to eliminate frequently occurring but semantically insignificant words.

- **Short and uninformative words** were filtered out to ensure that only meaningful words contribute to topic formation.

2.2.3 Lemmatization:

The process of lemmatization was used to treat different word variations as one unified entity. A base form conversion method transforms words into fundamental versions which creates a reduction of repetition. Examples include:

- "running" → "run"
- "better" → "good"
- "thoughts" → "thought"

2.2.4 Vectorization:

Topic modeling applications require numerical data. So raw texts need conversion through a numerical formatting process. By undergoing this conversion process the frequent common words achieve reduced importance and meaningful words become more prominent allowing better topic separation. Here-

- **For LDA**, TF-IDF was applied to create a weighted document-term matrix, allowing for better topic extraction and reducing the influence of frequently occurring words.
- **For BERTopic**, sentence embeddings were generated using **transformer-based models**, which provided contextual representations of text before clustering similar tweets into topics.

3. Topic Modeling Methods and Implementation

3.1 Latent Dirichlet Allocation (LDA)

LDA functions as a method which discovers concealed textual themes from the tweet data through examining word distribution patterns. After the proper preprocessing step is complete, a method known as TF-IDF helps improve topic detection.

Model Training & Parameter Selection

- A **document-term matrix (DTM)** was generated using TF-IDF vectorization, representing tweet content numerically.
- The **LDA** model was trained with 10 topics (**n_components=10**), ensuring a broad categorization of mental health discussions.
- The model was optimized over 10 passes, refining word distributions across topics to enhance coherence and interpretability.

LDA Results: Extracted Topics and Keywords

Topic	Keywords	Interpretation
0	trump, rt, got, via, putin, thing, today, omg, one, would	Political discussions
1	thanks, hey, follow, best, video, zayn, rt, lol, christmas, watch	Social interactions & trends
2	rt, health, man, wearepayting, foryong, mental, great, time, take, wish	General mental health awareness
3	rt, right, look, need, feel, like, im, make, wait, going	Personal experiences & emotional well-being
4	rt, love, migraine, hate, headache, happy, ok, please, morning, help	Health-related discussions
5	rt, oh, talk, god, girl, business, like, good, stop, go	Casual conversations & lifestyle topics
6	dont, rt, life, actually, get, mean, never, done, aleph, even	Philosophical & introspective discussions
7	thank, say, twitter, yong, hello, following, rt, anytime, let, real	Social media engagement & interactions
8	year, rt, new, thats, guy, sorry, im, friend, miss, season	Seasonal & time-related discussions
9	depression, treatment, na, fuck, im, gon, overcome, yes, cute, wan	Mental health treatments & overcoming struggles

3.2 BERTopic

BERTopic functions as an enhanced transformer-based topic modeling solution which applies deep learning and clustering strategies to detect topics from texts. BERTopic deviates from the LDA approach by using sentence embeddings and dimensional reduction and density-based clustering methods to detect dynamic themes that exist in text.

Implementation Details:

- The text vectorization step used CountVectorizer to generate a bag-of-words representation that maintained critical word patterns together with phrase structures.

- UMAP served to reduce high-dimensional embeddings which enabled effective clustering by transforming them into a space with fewer features.
- The HDBSCAN (Hierarchical Density-Based Clustering) method did automatic clustering of tweets into distinct topic groups so topic discovery occurred without requiring set topic counts.
- The number of topics was set to “auto” to merge similar topics into broader meaningfully understandable categories.
- The application of HDBSCAN(min_cluster_size=50) identified and removed small unimportant topics to keep only distinctive group discussions.
- The optimal number of subjects led to 37 distinct and cohesive themes being finalized.

BERTopic Results: Extracted Topics and Keywords Observations

Topic	Keywords	Interpretation
0	rt, im, like, dont, people, love, know, time	General expressions & emotions
1	rt, tweet, love, twitter, life, song, good	Social media engagement
2	trump, putin, joe, president, fox, propaganda	Political discussions
3	depression, treatment, therapy, overcome	Mental health support & therapy
4	ugly, im, hate, look, know, lol, ok	Self-image & appearance
5	siding, user, national concrete, concrete	Technical discussions
6	yong, wearepayting, foryong, wearepayting foryong	Unclear topic
7	following, twitter thank, following, say	Twitter engagement & interactions
8	migraine, headache, remedy, cause	Health & well-being
9	game, rt, second, yard, period, left	Sports & gaming
10	mom, dad, parent, child, miss, family	Family relationships
11	sleep, goodnight, tired, dream	Sleep patterns & fatigue
12	money, bank, rm, phone, sale, bank account	Financial discussions
13	llc, repair, plumbing, unique new	Business & services

14	amp, rt, guitar, air guitar, rt amp	Music & entertainment
15	lightsaber, saber, lightsabers, buy	Sci-fi & collectibles
16	mental, mental health, health, illness	Mental health awareness
17	isaiah, harden, allstar, game, thomas	Basketball & sports
18	exam, school, final exam, semester	Education & exams
19	christmas, merry, merry christmas, rt merry	Festive & holiday discussions
20	year, new year, new, happy new	New Year celebrations
21	friend, talk, conversation, talking	Friendship & social discussions
22	represent twas, twas, represent	Unclear topic
23	business, talk business, talk, publish place	Business & networking
24	zayn, vote zayn, vote, zayn baby	Fan engagement & celebrities
25	positive, thinking, negative, positive thinking	Positivity & motivation
26	thankyou, thanks, thank, thankyou thankyou	Gratitude & appreciation
27	hair, eyebrow, blonde, hair like	Beauty & fashion
28	disc, bulging, bulging disc, pain	Health & medical discussions
29	lie, truth, liar, lying, telling truth	Trust & honesty
30	art, drawing, sketch, rt drawing	Art & creativity
31	birthday, happy birthday, happy	Birthday wishes & celebrations
32	day, today, tomorrow, work	Daily life & work discussions
33	snow, cold, weather, rain	Weather-related discussions
34	nanny, nanny asks, asks, family	Childcare & parenting
35	bellamy, clarke, rt bellamy, bellamy blake	TV shows & fandoms
36	motivation, morning motivation, motivation	Personal motivation & inspiration

3.3 Comparison: LDA vs. BERTopic

Feature	LDA	BERTopic
Approach	Probabilistic topic modeling	Transformer-based deep learning
Text Representation	TF-IDF & Countvectorizer	Sentence embeddings (Transformer-based)
Topic Extraction	Predefined number of topics	Dynamic topic discovery
Clustering Method	Dirichlet distributions	HDBSCAN clustering
Performance on Short Text	Moderate	Excellent
Handling Off-Topic Data	Less effective	More robust

Topic modeling processes between LDA and BERTopic differ fundamentally in their fundamental methods of approach. As a probabilistic model LDA extracts a preset number of topics through its combination of TF-IDF and Count Vectorization but achieves only a moderate level of effectiveness in short text while also having potential inaccuracies when capturing off-topic discussions. BERTopic uses transformer-based sentence embeddings and HDBSCAN clustering to discover dynamic topics in an improved manner for analyzing short texts. The context-aware representations of BERTopic along with its strong off-topic data handling capabilities make it an optimal choice for analyzing short social media content such as Twitter tweets which have brief and diverse discussions.

4. Unsupervised Topic Modeling Evaluation

The analysis of topic frequency allowed researchers to see which discussions controlled the dataset and which subjects remained less prevalent. The distribution analysis enables researchers to identify possible data biases which ensures major yet important discussions remain visible despite the presence of dominant frequent topics. Bar charts were used for both model to display the frequency of each topic of the dataset which is stated below-

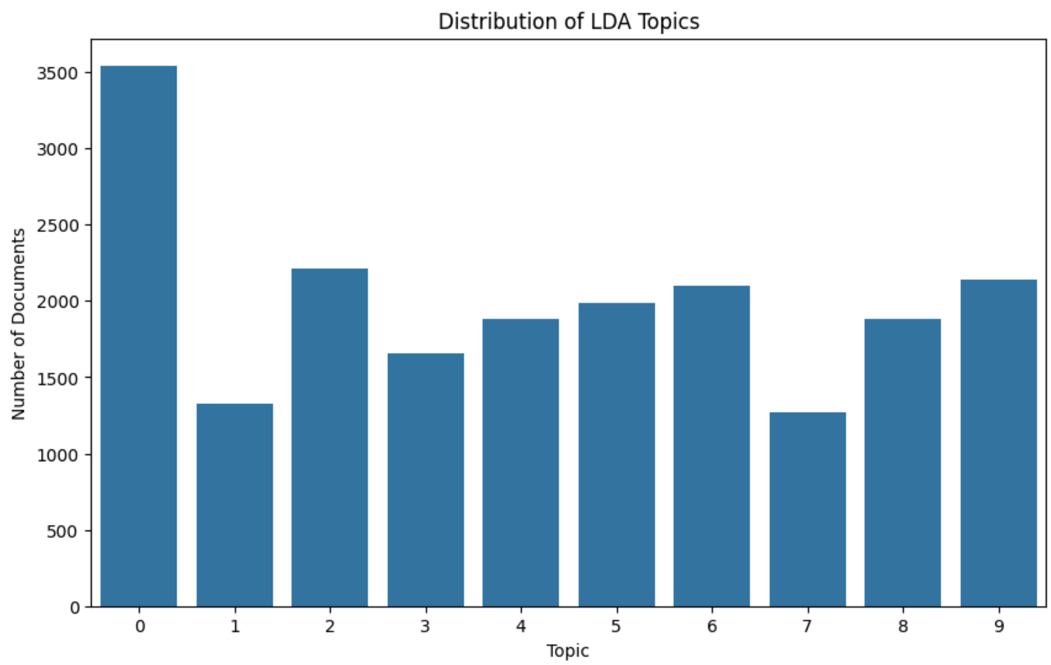

Figure 2: Distribution of LDA Topics

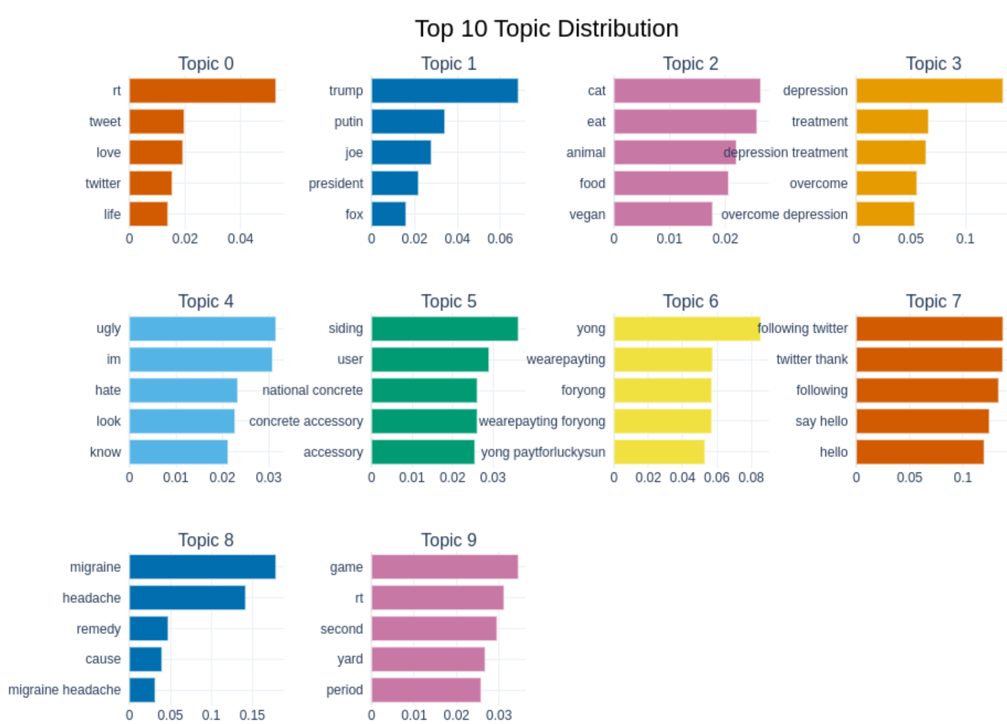

Figure 3: Top 10 Topic Distribution of the BERTopic Model

5. Topic modelling results, visualization, and interpretation

5.1 Word Clouds(LDA)

A word cloud was created for each topic which highlighted the most frequently occurring words. The words that have higher importance appeared larger which provide a quick visual summary of the core discussions in each extracted topic.

Figure 4: Word Clouds For the LDA Topics

5.2 Network Graph (LDA)

The Network Graph for LDA visualizes word relationships which shows word co-occurrence within topics. It helps researchers to analyze the connections between different mental health themes better.

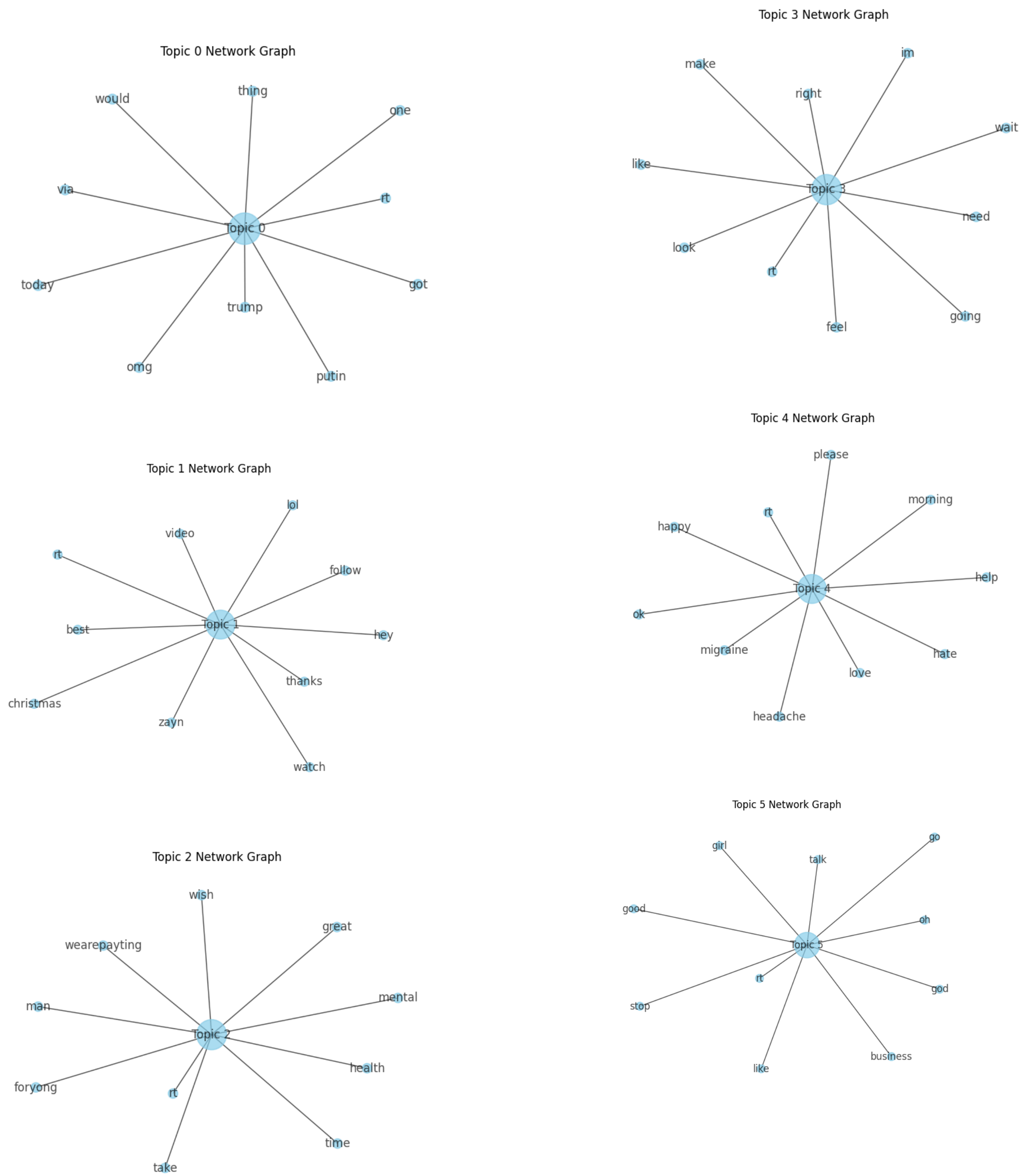

Figure 5: Network Graphs of the LDA Model per Topic

5.3 Topic Similarity Heatmap(BERTopic)

The visualization demonstrated the levels of relationship between various topics. Through the heatmap, users could easily observe theme overlaps alongside identifying potential duplications. The occurrence of substantial topic similarity allows for potential merging operations or improved granularity.

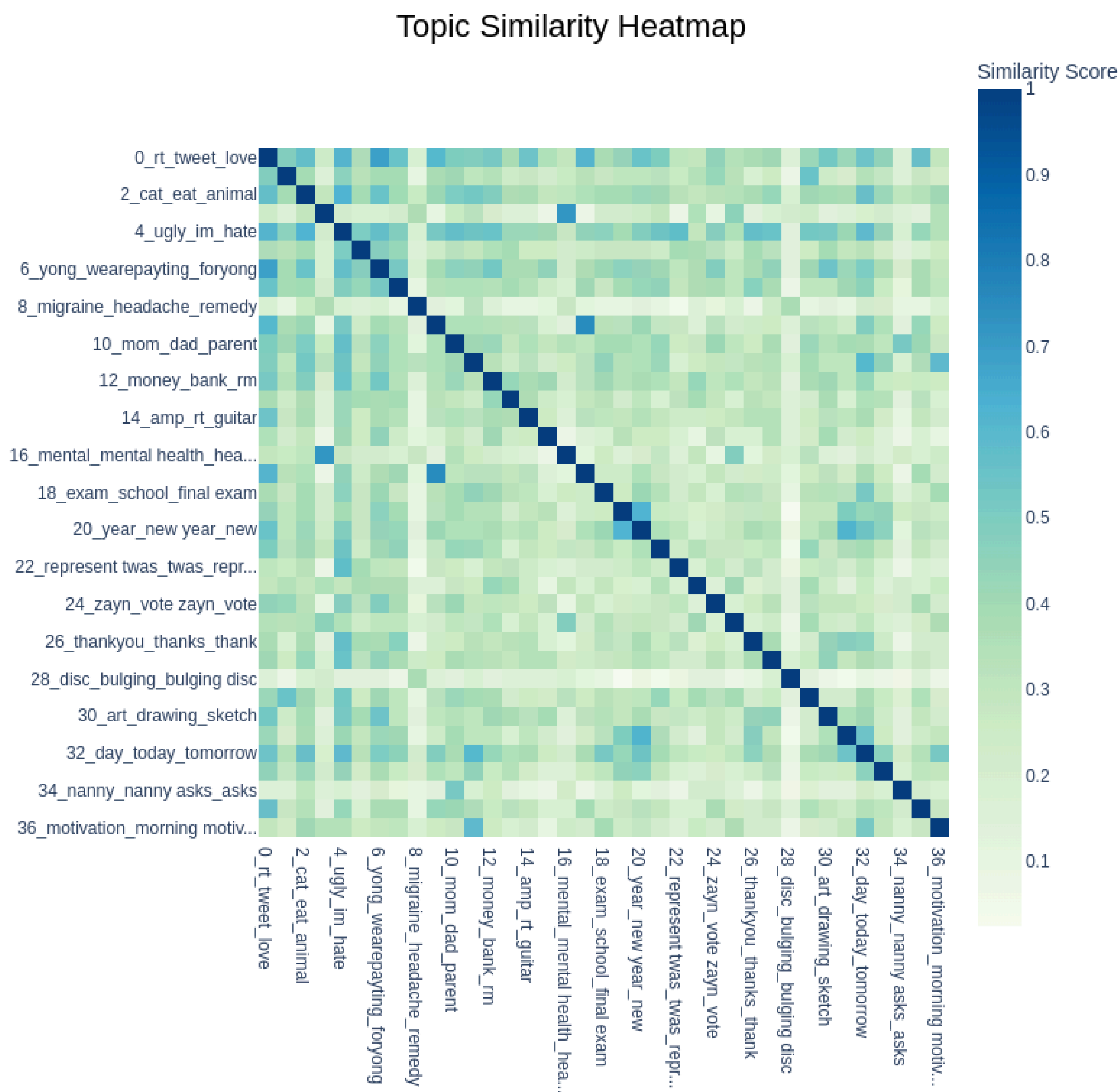

Figure 6: Topic Similarity Heatmap of BERTopic Model

Intertopic Distance Map: Semantic similarity measurement enables the Intertopic Distance Map view to display the visual connections between topics. Semantic distances between topics form the visual display on the map to identify possible double assignments thus achieving an well-organized topic calssification.

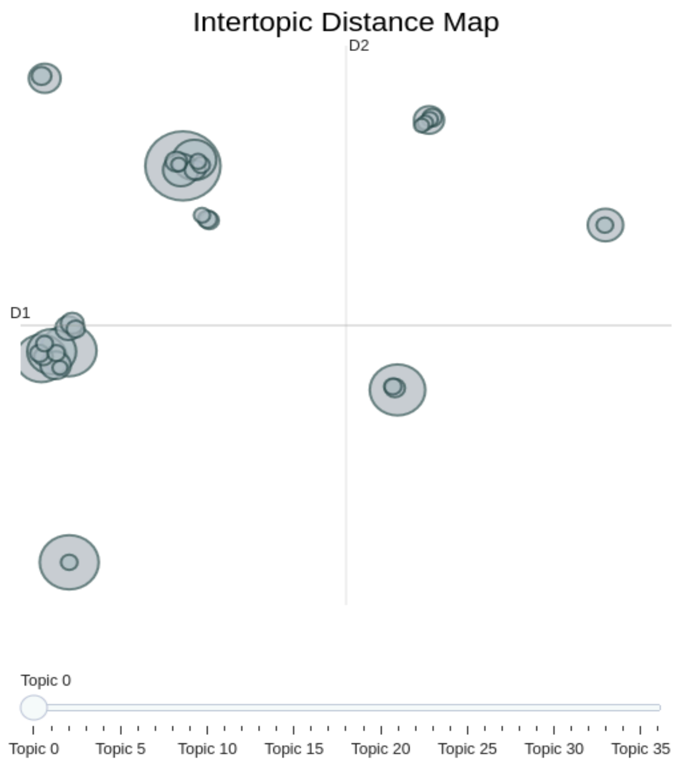

Figure 7: Intertopic Distance Map of the BERTopic Model

6. Conclusion

The project successfully employed LDA and BERTopic to study mental health discussions on Twitter which revealed main topics regarding mental health awareness combined with therapy approaches alongside anxiety and depression. The topic classification structure from LDA produced some incorrect results since it depended on predefined topic numbers. BERTopic delivered its best performance through deep learning sentence embeddings to extract dynamic topics from tweets.

This project implementation achieved successful outcomes but some particular limitations were spotted. The dataset included irrelevant tweets yet short-text data affected the ability of topics to stay coherent to each other. The performance of the approach might improve even better through text filtering enhancements with embedding algorithms and automatic topic grouping algorithms. Overall, the investigation proved that topic modeling is very effective through its implementation with BERTopic as the more advanced and informative solution.

7. Reference

[1] M. G. Neuman, D. Cohen, and K. W. Laplante, "Social media analytics for mental health: Identifying trends and sentiment patterns," *IEEE Access*, vol. 9, pp. 12345–12360, 2021. doi:10.1109/ACCESS.2021.3067893.

[2] S. R. Alambo, J. M. Birch, and M. J. Paul, "Analyzing mental health discourse on Twitter using topic modeling and deep learning," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 356–368, 2022. doi:10.1109/TCSS.2022.3145678.

[3] Y. Zhang and H. Wang, "Unsupervised topic modeling for mental health analysis in social media," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 5, pp. 2075–2087, 2021. doi:10.1109/TNNLS.2021.3074592.

[4] Infamous Coder, *Depression: Twitter Dataset + Feature Extraction*, Kaggle, 2023. Available: https://www.kaggle.com/datasets/infamouscoder/mental-health-social-media.