

<ANALYSING MENTAL HEALTH DISCOURSE ON MINDLINE.SG>

Johnathan Yap Wen Jie¹

¹Health and Wellbeing

john.yap@nus.edu.sg [Email of corresponding author only]

SUB-THEME

Opportunities from Wellbeing

KEYWORDS

mental health, sentiment analysis, web scraping, topic modelling

CATEGORY

Poster

ABSTRACT

BACKGROUND

In June 2020, Singapore launched mindline.sg, an anonymous digital mental health resource website. As part of its resources, the platform includes [let's talk](#), a community forum promoting open mental health discussions and offering resources, education, and support among local students.

OBJECTIVES

This poster aims to textually analyse discussions within mindline.sg's *let's talk* forum, focusing on identifying prevalent wellbeing themes through analysing sentiment patterns and topic modelling. By understanding these aspects, this poster hopes to derive insights that will inform strategies for promoting mental wellness and enhancing psychological safety among students.

METHODS

Data Collection and Preprocessing

let's talk was selected as the primary data source for this poster due to its role as a public forum for mental health discussions among local students.

A total of 1,086 unique forum posts (from the 'Ask a Therapist' section) were systematically collected using web scraping techniques with Python (i.e., BeautifulSoup and Selenium), covering the period from June 2022 to June 2024.

Pre-processing text normalisation techniques such as lowercasing, stop words removal and tokenisation were applied to standardise the textual data for further analysis.

Sentiment Analysis

Sentiment analysis was conducted to assess the emotional tone of these forum discussions. The analysis utilised the VADER (Valence Aware Dictionary and Sentiment Reasoner) sentiment analysis tool provided by NLTK (Natural Language Toolkit). The VADER sentiment analyser computed sentiment scores for each post, generating a compound score that reflects overall sentiment polarity (positive, neutral, or negative).

Posts were categorised based on their compound sentiment score:

- Positive: Posts with a compound score ≥ 0.05 .
- Negative: Posts with a compound score ≤ -0.05 .
- Neutral: Posts with compound scores between -0.05 and 0.05

Topic Modelling

To uncover the primary themes discussed in the forum posts, this poster utilised Latent Dirichlet Allocation (LDA), a widely used probabilistic model in natural language processing. LDA identifies clusters of words that frequently occur together across posts, assuming these clusters represent coherent topics.

RESULTS

Sentiment analysis (see Figure 1) revealed a range of emotional tones in forum discussions, including positive, neutral, and negative sentiments. The prevalence of negative sentiments, which was 21% higher than positive and 59% higher than neutral, suggests that the anonymity of the *let's talk* forum provides a safe space for participants to openly express their challenges and emotional distress. This anonymity likely fosters a sense of psychological safety, enabling students to discuss sensitive topics without fear of judgment or repercussion.

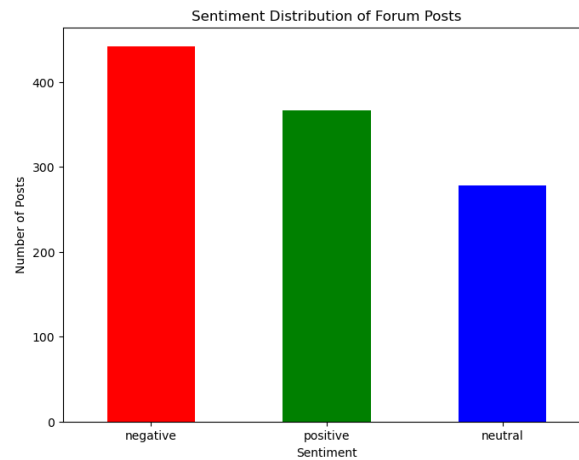


Figure 1. Sentiment Analysis of Forum Posts

The temporal analysis (see Figure 2) revealed spikes in negative sentiments, particularly in January 2024 (i.e., start of a new school term) and mid-April 2024 (i.e., exam periods). These spikes suggest heightened stress and challenges faced by students during these periods, which are critical for informing targeted interventions and support strategies.

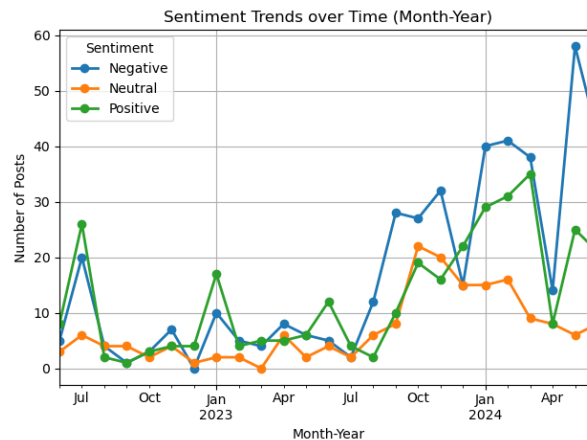


Figure 2. Temporal Sentiment Analysis

Using Latent Dirichlet Allocation (LDA), five key themes were identified in forum discussions:

- **Personal Experiences and Emotions:** Topics included discussions about work, social interactions, and daily life.
- **Emotional Expression and Relationships:** Topics centered on feelings of loneliness, social connections, and personal identity.
- **Existential Reflections and Growth:** Discussions focused on life reflections, personal development, and aspirations.
- **Family Dynamics and Personal Challenges:** Themes explored familial relationships, school experiences, and personal responsibilities.
- **Seeking Advice and Support:** Posts often sought advice, support, and shared personal challenges.

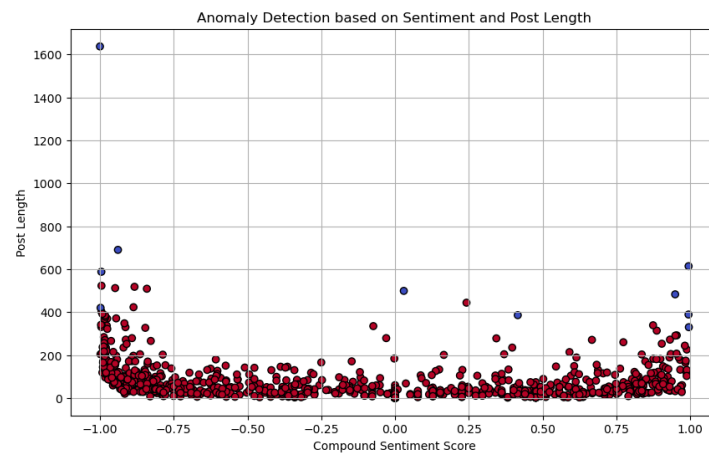


Figure 3. Isolation Forest (anomaly detection).

An Isolation Forest algorithm (see Figure 3) was employed to detect anomalous posts within the dataset. By analysing features such as sentiment scores, post length, and content structure, it flagged posts with unusual patterns, suggesting potential extreme emotional responses that could indicate crisis situations requiring further attention.

In total, 10 posts were flagged as anomalous, representing 1.1% of the total posts. One notable example included a detailed account from a student describing verbal and physical abuse from her mother, underscoring the serious nature of some flagged content.

CONCLUSION AND FUTURE DIRECTIONS

This poster examined discussions from mindline.sg's *let's talk* forum using sentiment analysis and topic modelling, highlighting prevalent negative sentiments and key thematic clusters among students. Future directions could explore the dynamics and interactions between different users within the forum, leveraging network analysis to understand community support structures and identify influential contributors. These insights could further inform intervention strategies and enhance the efficacy of mental health support initiatives in educational contexts.

REFERENCES

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.
- Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth ieee international conference on data mining* (pp. 413-422). IEEE.
- Weng, J. H., Hu, Y., Heaukulani, C., Tan, C., Chang, J. K., Phang, Y. S., ... & Morris, R. J. (2024). Mental Wellness Self-Care in Singapore With mindline. sg: A Tutorial on the Development of a Digital Mental Health Platform for Behavior Change. *Journal of Medical Internet Research*, 26, e44443.

RECOMMENDED CITATION

[Citation in APA format starts here (**For Official Use Only**):

Author1 & Author2 (2024). Title of extended abstract [Paper presentation/Lightning talk/Poster presentation]. In *Higher Education Conference in Singapore (HECS) 2024, 3 December, National University of Singapore*. (URL of extended abstract)]