Asian Journal of Psychiatry 92 (2024) 103876

Contents lists available at ScienceDirect

Asian Journal of Psychiatry

journal homepage: www.elsevier.com/locate/ajp

Mental disturbance impacting wellness dimensions: Resources and open research directions

Muskan Garg

Mayo Clinic, Rochester, 55901 MN, USA

ARTICLE INFO

Keywords:
Depression
Mental health
Suicide risk
Wellness dimensions

ABSTRACT

In light of the unparalleled pressure faced by the healthcare system, there arises a pressing need for innovative solutions to comprehensively assess the overall well-being of individuals affected by mental health issues. With the objective of advancing AI-driven mental health analysis towards fine-grained analysis, we develop and publicly release our datasets, MULTIWD and WELLXPLAIN, specifically designed to capture the impact of mental disturbances on wellness dimensions in self-narrated texts. To this end, we make two major contributions. First, our examination focuses on the identification of one or more of the six distinct wellness dimensions evident within a given text, shedding light on the significant ramifications of mental disturbance, which, in turn, can perpetuate further mental unrest. Second, we conducting an extensive analysis of the textual cues that signify the presence of various wellness dimensions. We delve into the content of the text, examining specific linguistic and contextual markers that provide indications of the wellness dimensions being discussed. Finally, we open up future research directions to facilitate advancements in the domain of AI-driven approaches for fine-grained mental health analysis. This framework aims to establish and validate new clinical categories for mental distress, bridging the gap between mental wellness and illness, in response to the higher prevalence of distress compared to illnesses.

1. Introduction

The findings of a major epidemiological study in India indicate that approximately one in seven individuals have experienced a mental disorder in their lifetime, and the contribution of mental disorders to the total disease burden in India nearly doubled between 1990 and 2017 Sagar et al. (2020). Despite an estimated prevalence of approximately 6% for depression and 4% for anxiety disorders among the Indian population [1], the availability of services to meet the needs of the population remains limited, leading to a substantial treatment gap of 85% for common mental disorders Gururaj et al. (2016). Reports released in August 2021 Read et al. (2018) indicate that *1.6 million people* in England were on waiting lists to seek professional help with mental health care. In the US, mental health crises overwhelm the limited population of mental health professionals (MHPs), with 60 Million patient visits to primary care and 6 Million emergency visits annually. In a world grappling with an insatiable demand for quality healthcare in clinical psychology, the supply has struggled to keep pace, exacerbated by the recent pandemic. Such an overwhelming rise in the number of patients as compared to MHPs necessitated the use of social media platforms to express their thoughts with ease.

In parallel, the field of artificial intelligence (AI) has been catalyzing a revolution in the realm of medicine and healthcare Sun et al. (2023). There exists an apparent imbalance in the use of AI applications in mental health research, predominantly focus on studying depressive and other psychotic disorders, highlighting a significant gap in our understanding the broader spectrum of overall well-being. There has been a growing recognition of the need to better understand and address the spectrum of mental health that lies between the poles of wellness and illness. While mental illnesses have long been the focus of clinical attention and research, there is an increasing awareness of a significant portion of the population who experience mental distress that does not necessarily fit into traditional diagnostic categories. This has led to efforts aimed at defining and validating new clinical categories, which aim to encompass these experiences of distress. Such critical endeavors acknowledge the need to provide appropriate support and resources to a larger segment of the population, many of whom may not meet the criteria for a mental illness yet still struggle with significant challenges in their daily lives.

Our society embrace the concept of wellness not merely as the lack of

E-mail addresses: garg.muskan@mayo.edu, muskanphd@gmail.com.

https://doi.org/10.1016/j.ajp.2023.103876
Received 24 June 2023; Received in revised form 12 December 2023; Accepted 15 December 2023
Available online 20 December 2023
1876-2018/

disease, illness, and stress, but as the existence of a positive purpose in life, gratifying work and recreation, blissful relationships, a sound physical well-being, a nurturing environment, and overall happiness. This aligns with the foundational six wellness dimensions put forth by Halbert L. Dunn to provide a framework for understanding and cultivating overall well-being for a more balanced and holistic approach to wellness. The six wellness dimensions are physical aspect (PA), emotional aspect (EA), intellectual aspect (IA), social aspect (SA), vocational aspect (VA), and spiritual aspect (SpA).

In-person clinical psychologists observe a decline in wellness among individuals resulting from factors such as social isolation, inadequate hospitalization, and impaired performance in daily activities. These circumstances indicate the potential impact on various dimensions of wellness as a result of underlying causes contributing to mental distress Parry et al. (2022). As such, the MHPs has recognized the imperative of analyzing firsthand voluntary written texts, given the prevalent inclination of individuals to express their thoughts more readily on (i) public health forums (e.g., dialogue4health), (ii) online communities (e.g., r/depression subreddit on Reddit), (iii) Talklife Kruzan (2019), rather than during in-person sessions. MHPs perceive social media profiles as a reliable source for conducting mental health analysis Iacus and Porro (2021), Yeung (2018).

The use of publicly available language resources is a well-established method employed by academic researchers and mental health practitioners to analyze human-writings and other self-narrated texts for studying human behavior Garg (2023). Garg et al. (2022) reveals the presence of two primary underlying factors contributing to mental disturbances, namely, (i) relationships and (ii) alienation, as evidenced by an analysis of the natural distribution of 5051 Reddit posts Garg et al. (2022). To this end, by thoroughly investigating the impact of mental disturbance (cause) on an individual's holistic well-being (consequences), we aim to establish a comprehensive framework for the integration of computational approaches in *mental disturbance impacting wellness dimensions*.

2. Methods

Consider the example post *P*:

P: `I'm 21 years old. I have aspergers syndrome and depression ← (Physical Aspect), I have struggled quite a lot and I want to do stuff my own way to get better (with the help from actual professionals). My mum, dad and step-mum ← (Social Aspect) won't leave me alone and they constantly make choices for me and it's starting to get to me. They make me feel unhappy and miserable ← (Emotional Aspect). What should I do?`

In a post P, a user who is 21 years old has referenced their Asperger's syndrome and depression, indicating a medical issue that significantly affects their physical wellbeing. They also describe issues with their interpersonal relationships, specifically with their mother, father, and step-mother, who seemingly infringe upon their personal space, affecting their social wellness. Moreover, the user expresses feelings of unhappiness and misery due to their parents' intrusion, which impacts their emotional wellbeing. By taking into account these multifaceted dimensions of wellness reflected in this post, we can develop AI models that are better equipped to identify and emphasize particular aspects of deteriorating mental health.

With the goal of facilitating the research community with publicly available language resources in exploring the interconnections within mental health and Wellness Dimensions (see Fig. 1), we have constructed and released two datasets:

1. MULTIWD,[1] a six-label text classification dataset, devised for discerning one or more wellness dimensions present in a given text.
2. WELLXPLAIN,[2] a multi-class classification dataset designed specifically to determine one of the predefined Wellness Dimensions in a given text along with the textual spans signifying its presence.

2.1. Data acquisition

2.1.1. MULTIWD

A total of 4000 instances were collected from the subreddits `r/depression` and `r/SuicideWatch` during the period of November 23, 2021, to January 4, 2022. To ensure dataset *diversity*, an average of approximately 100 data points per day was collected. The data points were manually cleaned by removing all the empty posts, the posts containing only URLs or other social media handles, advertisements and irrelevant data. Additionally, in order to maintain consistency among the data points and with future AI-driven mental health analysis in mind, we imposed a length restriction on each post, ensuring that they were limited to a maximum of 300 tokens. We finally ensure that our dataset only contains relevant and meaningful data, resulting in 3281 samples. We annotate each post to identify the presence of one or more of the six wellness dimensions within the text.

2.1.2. WELLXPLAIN

To analyze the wellness dimensions within a text, we break 1000 samples into smaller sections from MULTIWD dataset. This approach allows us to examine each section individually and identify the specific wellness dimensions present within them. By employing this methodology, we can determine the textual cues that indicate the presence of one of the four primary wellness dimensions by introducing WELLXPLAIN dataset. The motivation behind reducing the initial six dimensions to four stems from the substantial overlap observed during inter-annotator agreement assessments for annotations related to (i) the Intellectual and Vocational aspect, and (ii) the Emotional and Spiritual aspect. This adjustment allows for a more focused and accurate assessment of the wellness dimensions present in the text, considering the complexities and intricacies of their interconnections. We performed expert-guided manual cleaning and filtered the Reddit posts to limit them to *user experience* impacting mental disturbance, resulting in a new English dataset of 3092 instances with 72,813 words.

2.2. Annotation scheme

Our team of experts developed comprehensive annotation guidelines that strike a balance between utilizing text-based markers for advanced AI models and employing psychological insights gained from reading between the lines, all based on Dunn's definitions Dunn (1959). These guidelines aim to achieve multiple goals: (i) ensuring accurate identification of wellness dimensions in alignment with Dunn's definitions, (ii) maintaining consistency among annotators with minimal errors or discrepancies, and (iii) enabling efficient annotation of a large amount of data within a reasonable timeframe. The guidelines provide detailed instructions, including specific examples and criteria for each dimension, as well as guidance on resolving ambiguous cases and ensuring overall consistency and accuracy throughout the annotation process.

2.3. Annotation task

We enlisted the assistance of three student interns enrolled in a postgraduate course to contribute to our annotation process. Our team of experts meticulously trained these interns using a well-constructed

[1] https://github.com/drmuskangarg/MultiWD
[2] https://github.com/drmuskangarg/WellnessDimensions/

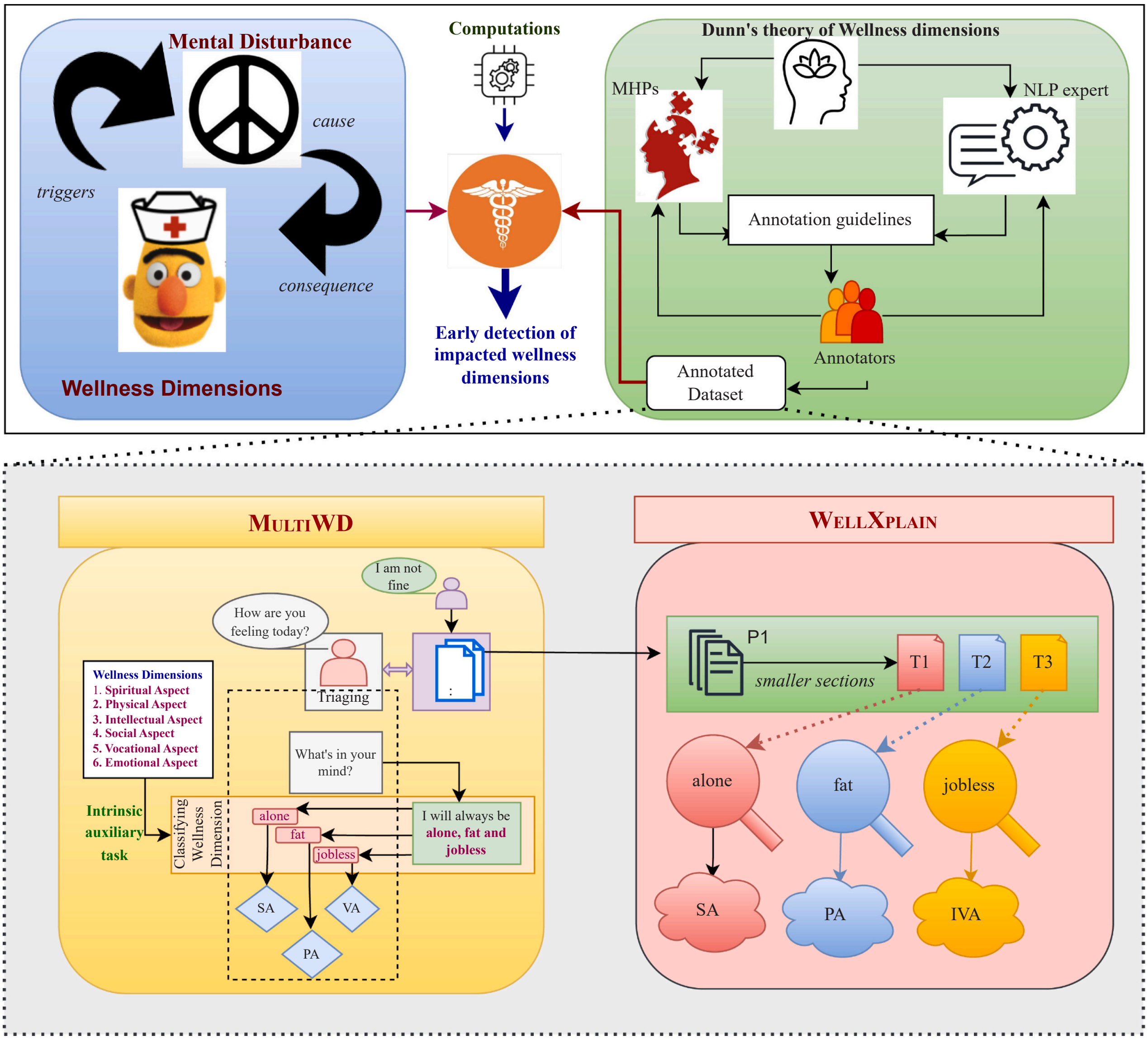

Fig. 1. Advancing the fusion of computational intelligence and clinical psychology to simulate the mental disturbance that affect mental disturbance and vice-versa. Our approach involves data collection, annotation, and public release, enabling the simulation of deteriorating mental health impacting wellness dimensions as evident from self-narrated texts. We convert a given post (P1) in smaller sections (T1, T2, T3, …) for fine-grained analysis.

annotation scheme, ensuring they were equipped with the necessary skills and understanding to accurately annotate the data.

For MULTIWD, we ask all three students to perform six labels demonstrating the presence (1) or absence (0) of every wellness dimension in a given text. The final annotations were verified by a senior clinical psychologist leveraging the annotation guidelines in-case of any discrepancies. Inter-annotator agreement (IAA) within the annotated dataset was assessed using Fleiss' Kappa coefficient.[3] The agreement on labels for each of the six aspects (i.e., SPA, PA, IA, SA, VA, EA) was examined individually, yielding kappa coefficients of 58.34%, 68.66%, 75.29%, 83.14%, 76.77%, and 64.23% respectively. The average agreement across all aspects was found to be 71.07%, indicating substantial agreement and ensuring *reliability*. In the context of a sensitive, domain-specific task involving psychology and emotion-based multi-labeling, it is common for inter-annotator agreement to be relatively low Roccabruna et al. (2022).

For WELLXPLAIN, we ask students to perform a two-fold annotation for each short-text: (i) Assignment of Wellness label (individual annotations) and (ii) Identifying text spans that explain the reason for the given label (group annotations). After, experts' validation on final annotations, we perform IAA for both the tasks, resulting in a kappa score of 74.39% and 87.32%, respectively. To ensure consistency, the final annotations were determined through a majority voting mechanism, followed by additional verification by experts. While there was slightly lower inter-annotator agreement in marking the classes of wellness dimensions, particularly between (Physical Aspect) PA and (Spiritual and Emotional Aspect) SEA, a higher agreement was observed for the selection of explanatory text spans.

[3] https://en.wikipedia.org/wiki/Fleiss%27_kappa

3. Results and discussions

3.1. MULTIWD: data analyses

Within the context of the MULTIWD dataset, it is evident that there exists a notable degree of variability in the length of each individual post, ranging from 6 to 300 tokens. Posts attributed to different dimensions exhibit an average token count spanning between 122 and 155, showcasing the diverse nature of textual content within distinct wellness dimensions. Furthermore, Based on our analysis, where we have counted the occurrence of each label in our corpus of 3281 posts, we have identified a total of 6114 labels distributed across the six wellness dimensions. The natural distribution of these wellness dimensions is as follows: SPA (Spiritual Wellness) with 200 occurrences (6.09%), PA (Physical Wellness) with 923 occurrences (28.13%), IA (Intellectual Wellness) with 651 occurrences (19.84%), SA (Social Wellness) with 2129 occurrences (64.88%), VA (Vocational Wellness) with 550 occurrences (16.76%), and EA (Emotional Wellness) with 1661 occurrences (50.62%) (see Table 1). The variations in occurrence across the different dimensions shed light on the diverse nature of the impacts of mental disturbance on various aspects of wellness, providing insights for further research and potential targeted interventions. Additionally, this distribution highlights SPA and SA as the least and most impacted wellness dimensions, respectively, when considering the influence of mental disturbance in social media posts Gireesan (2022). Consequently, the discussions held on social media platforms have a pronounced influence on interpersonal relationships, notably affecting social and emotional well-being more prominently than other dimensions such as intellectual or vocational aspects.

While our findings provide substantial insights into the wellness dimensions impacted by mental disturbance, particularly in relation to interpersonal relations and social media discussions, it is important to acknowledge that alternative platforms may offer more comprehensive information regarding physical well-being (such as electronic health records with hospital visits) or intellectual/vocational aspects (such as interactions on LinkedIn). Moreover, the average number of sentences was found to be 8, indicating the incorporation of multiple wellness dimensions within the entirety of the text. This observation suggests that the texts encompass a diverse range of wellness dimensions, highlighting the multidimensional nature of the discussions or narratives present in the dataset.

3.2. WELLXPLAIN: text analysis

For in-depth analysis we find textual cues suggesting one of the four wellness dimensions in a short text, usually comprising of a single sentence. We notice the most frequently occurring words as: (i) Physical Aspect: ugly, anxiety, sleep, meds, pain, tired, panic, drunk, alcohol, diagnosed; (ii) Intellectual and Vocational Aspect: job, school, work, college, money, failing, failed, time, life, year; (iii) ocial Aspect: friends, alone, family, lonely, people, feel, want, someone, parents, friend; and (iv) Spiritual and Emotional aspect: hate, feel, sad, worthless, motivation, anxiety, life, cry, shit, useless.

The lists pertaining to each aspect exhibit notable dissimilarities, with only a handful of shared words, such as "feel," present in both the SA and SEA lists. For instance, the expression "feeling lonely" was assigned to the SA category, whereas "feeling useless" was assigned to the SEA category. This observation underscores the fact that while certain words may possess associations with multiple aspects, their classification is heavily influenced by the specific context and connotations surrounding them.

3.3. Open directions

We open up the following set of research questions to open new research directions:

1. Research Question 1 (RQ1): Is it feasible to construct AI-driven multi-label text classification models that are both safe and explainable, enabling the identification of one or more wellness dimensions within a given text?
2. Research Question 2 (RQ2): Can we develop an explainable AI-driven multi-class classifier that prioritizes paying more attention to textual spans indicating a specific wellness dimension within a given text, thus enhancing the likelihood of accurate classification?
3. Research Question 3 (RQ3): Can we successfully develop a safe and explainable AI-driven model trained on individual aspect identification within the WELLXPLAIN dataset, and subsequently apply it to the multi-dimensional task of classifying wellness dimensions within the MULTIWD dataset?
4. Research Question 4 (RQ4): Is it feasible to create a discourse analysis or text-spans extraction mechanism to detect textual indicators indicating wellness dimensions?

In pursuit of this objective, we anticipate significant progress in the realm of deep learning and transfer learning across two datasets with varying characteristics, encompassing overall well-being. Researchers continue to explore the potential of Large Language Models like ChatGPT, LLAMA, and Med-Palm in the field of psychiatry Thornton et al. (2023), a question that remains an open and evolving area of study.

3.4. Ethics, limitations and broader impact

3.4.1. Ethics

To ensure privacy protection and prevent misuse, all sample posts have been anonymized, obfuscated, and rephrased. Our approach strictly adheres to privacy regulations, ensuring the non-disclosure of any personal information. In our work, trustworthiness plays a crucial role. It pertains to ensuring the reliability, credibility, and ethical integrity of my research methodology, data analysis, and findings. By upholding trustworthiness, we strive to conduct my research with transparency, accuracy, and adherence to ethical guidelines. This involves employing rigorous data collection methods, employing systematic and transparent data analysis techniques, and addressing potential biases or limitations. By prioritizing trustworthiness, we aim to enhance the credibility and impact of my research, fostering confidence in the validity and applicability of my findings among other researchers and practitioners in the field. Our ethical practices shall enable the research community to develop safe and explainable AI-driven models through fine-grained mental health analysis. We adhere to FAIR principles and

Table 1
The statistics of MULTIWD for different dimensions of wellness, indicating the imbalanced nature of dataset.

Dimensions	Count	Number of Words				Number of Sentences		
		Min/ post	Avg/ post	Max/ post	Total	Avg/ post	Max/ post	Total
Spiritual Aspect	200	6	122.625	298	24,525	8.145	42	1629
Physical Aspect	923	8	134.678	300	124,308	8.55	32	7892
Intellectual Aspect	651	15	145.285	300	94,581	9.01	31	5869
Social Aspect	2129	1	132.728	300	282,580	8.29	42	17,652
Vocational Aspect	550	13	155.843	300	85,714	9.814	30	5398
Emotional Aspect	1661	1	131.414	300	218,280	8.279	31	13,752

make our dataset Findable, Accountable, Interoperable, and Reliable.

3.4.2. Limitations

Given the nature of Reddit as a community-based discussion platform, we make the assumption that the user-generated posts on Reddit reflect the authentic expression of users' mindsets and thought processes. Unlike content that may be tailored for attention-seeking purposes on other social media platforms, we believe that the discussions taking place within the subreddits `r/depression` and `r/SuicideWatch` are genuine interactions initiated by individuals experiencing mental distress. This assumption is based on the understanding that these individuals seek support, understanding, and connection within these specific subreddit communities, contributing to a more authentic representation of their experiences and perspectives. As a starting point of our study, we assume that while infusing any textual cues indicating wellness dimension in subreddits `r/depression` and `r/SuicideWatch`, the Reddit user undergo psychological disturbance.

3.4.3. Broader impact

Our research aims to provide a solid foundation for the integration of computational studies in real-time clinical psychology and psychiatry settings. By examining the impact of mental disturbance on wellness dimensions and analyzing textual cues, our work contributes to the development of computational approaches that can be applied in crucial areas like mental health triaging Wood and Anderson (2023), clinical diagnostic interviewing Schneider et al. (2022), and motivational interviewing Dalrymple et al. (2022). This research aims to leverage computational tools and techniques to improve and facilitate mental health assessments and interventions in real-time clinical practice. By identifying the wellness dimensions influenced by mental disturbance, it becomes possible to detect early signs of declining well-being during the onset of depression. This recognition and monitoring of specific impacted wellness dimensions enable timely interventions to address potential declines in overall well-being.

4. Conclusion

We conclude our work as the first of its kind, to examine multiple wellness dimensions impacted by mental disturbance in firsthand user-generated Reddit posts. We observe highly imbalanced nature of multi-label text classification dataset (MULTIWD) as evident by the count of the most impacted dimension (SA) which is about 10 times highers than the least impacted wellness dimension (SPA) identified in texts. Through our examination of textual cues, we gain valuable insights into how wellness dimensions are represented and expressed in texts, leading us to identify the most commonly used words. Ultimately, our research contributes to the problem domain of identifying wellness dimensions in social media posts for higher-level analysis using computational approaches. Furthermore, our research highlights the critical need for establishing new clinical categories to accurately capture the wide spectrum of mental distress, recognizing its prevalence and impact beyond the traditional confines of diagnosed mental illnesses, thereby posing a foundation for future research.

Financial disclosure

There is no financial disclosure required for this document.

CRediT authorship contribution statement

Muskan Garg: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

There is no conflict of interest.

Acknowledgment

We would like to sincerely acknowledge the valuable contributions of our postgraduate student annotators, Ritika Bhardwaj, Astha Jain, and Amrit Chadha, for their dedicated efforts in the annotation process. Our heartfelt gratitude goes to Veena Krishnan, a senior clinical psychologist, and Ruchi Joshi, a rehabilitation counselor, for their unwavering support throughout the project. We would also like to express our deep appreciation to Prof. Sunghwan Sohn for his continuous guidance and support.

References

Dalrymple, J., Dimas, K., Illes, R.A., Spradling, T., 2022. Health promotion and wellness. Family Medicine: Principles and Practice. Springer, pp. 95–106.
Dunn, H.L., 1959. High-level wellness for man and society. Am. J. Public Health Nations Health 49, 786–792.
Garg, M., 2023. Mental health analysis in social media posts: a survey. Arch. Comput. Methods Eng. 1–24.
Garg, M., Saxena, C., Saha, S., Krishnan, V., Joshi, R., Mago, V., 2022. Cams: An annotated corpus for causal analysis of mental health issues in social media posts. In: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 6387–6396.
Gireesan, A., 2022. Evolution of belongingness: its past, present, and future. Handbook of Health and Well-Being: Challenges, Strategies and Future Trends. Springer, pp. 97–127.
Gururaj, G., Varghese, M., Benegal, V., Rao, G.N., Pathak, K., Singh, L., Mehta, R., Ram, D., Shibukumar, T., Kokane, A., 2016. National mental health survey of india, 2015–16. Bengaluru, India: National Institute of Mental Health and Neuro Sciences.
Iacus, S.M., Porro, G., 2021. Subjective Well-Being and Social Media. CRC Press.
Kruzan, K.P., 2019. Self-injury Support Online: Exploring Use of the Mobile Peer Support Application TalkLife. Cornell University.
Parry, D.A., Fisher, J.T., Mieczkowski, H., Sewall, C.J., Davidson, B.I., 2022. Social media and well-being: a methodological perspective. Curr. Opin. Psychol. 45, 101285.
Read, J., Harper, D., Tucker, I., Kennedy, A., 2018. Do adult mental health services identify child abuse and neglect? A systematic review. Int. J. Ment. Health Nurs. 27, 7–19.
Roccabruna, G., Azzolin, S., Riccardi, G., 2022. Multi-source multi-domain sentiment analysis with bert-based models. In: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 581–589.
Sagar, R., Dandona, R., Gururaj, G., Dhaliwal, R., Singh, A., Ferrari, A., Dua, T., Ganguli, A., Varghese, M., Chakma, J.K., et al., 2020. The burden of mental disorders across the states of india: the global burden of disease study 1990–2017. Lancet Psychiatry 7, 148–161.
Schneider, L.H., Pawluk, E.J., Milosevic, I., Shnaider, P., Rowa, K., Antony, M.M., Musielak, N., McCabe, R.E., 2022. The diagnostic assessment research tool in action: a preliminary evaluation of a semistructured diagnostic interview for dsm-5 disorders. Psychol. Assess. 34, 21.
Sun, J., Dong, Q.X., Wang, S.W., Zheng, Y.B., Liu, X.X., Lu, T.S., Yuan, K., Shi, J., Hu, B., Lu, L., et al., 2023. Artificial intelligence in psychiatry research, diagnosis, and therapy. Asian J. Psychiatry, 103705.
Thornton, J., D'Souza, R., Tandon, R., 2023. Artificial intelligence and psychiatry research and practice. Asian J. Psychiatry, 103509.
Wood, J.D., Anderson, E., 2023. Triaging mental health emergencies: lessons from philadelphia. Law Contemp. Probl. 86, 29–53.
Yeung, D., 2018. Social media as a catalyst for policy action and social change for health and well-being. J. Med. Internet Res. 20, e94.