

Transforming women's mental health with large language models: Opportunities and challenges in digital interventions

Usha Rana¹ & Rupender Singh²

¹Indira Gandhi National Open University (IGNOU), New Delhi, India

²Khalifa University, Abu Dhabi, UAE

received: 28 August 2024;

revised: 25 September 2024;

accepted: 12 September 2024

* * * * *

Large Language Models (LLMs), such as ChatGPT, have recently emerged as significant tools in various sectors, including mental health. These AI-driven technologies offer new ways to support individuals by providing information, advice, and emotional comfort. For women's mental health, particularly in India, these technologies present both promising opportunities and serious challenges. The unique mental health issues faced by women, influenced by social, cultural, and economic factors, require tailored interventions. LLMs could help bridge the gap in mental health services, especially in rural and underserved areas, where access to traditional mental health care is often limited. However, the implementation of LLMs in this context is challenging. This article will explore these challenges and opportunities in detail, supported by relevant literature.

CHALLENGES IN IMPLEMENTING LARGE LANGUAGE MODELS FOR WOMEN'S MENTAL HEALTH

Bias and stereotyping

One of the most critical challenges in using LLMs for women's mental health is the risk of reinforcing existing biases and stereotypes. LLMs are trained on vast datasets that include texts from various sources, including books, articles, and social media posts. If these datasets contain biased or sexist content, the LLMs may internalize and reproduce these biases in their interactions. For instance, an LLM might unintentionally perpetuate harmful stereotypes about women's roles in society or provide advice that is not sensitive to gender-specific issues. Research has shown that AI systems, including LLMs, can inherit and amplify biases in their training data (Bolukbasi et al.,

2016). This issue is particularly concerning in the context of mental health, where biased advice could exacerbate existing psychological distress rather than alleviate it. Efforts to mitigate these biases require careful curation of training data and ongoing monitoring and adjustment of the models to ensure they provide equitable and unbiased support (Buolamwini & Gebru, 2018).

Digital divide and accessibility

The digital divide poses a significant barrier to the effective use of LLMs for mental health support, especially in rural and underserved areas of India (Rana & Singh, 2023a). Many women in these regions lack access to the internet, smartphones, or computers necessary to interact with LLM-based services. Even when these technologies are available, digital literacy remains a significant challenge. A large portion of the female population in rural India is unfamiliar with digital tools and platforms, which limits their ability to benefit from AI-driven mental health interventions (Balcombe et al., 2021).

Moreover, language barriers further complicate accessibility. While many LLMs are developed primarily in English, many women in India speak regional languages or dialects. The lack of multilingual support in many LLMs can exclude these women from accessing the mental health resources they need. Addressing these challenges requires expanding digital infrastructure and developing LLMs that can operate effectively in multiple languages and are designed with the local cultural context in mind.

Privacy and security concerns

Privacy and security are paramount when it comes to mental health support. Many women may feel uncomfortable sharing personal and sensitive information with an AI, especially if they are still determining how it will

be used or stored (Rana & Singh, 2023b). This concern is heightened in India, where data privacy laws are still evolving, and there is limited public awareness about digital privacy rights.

LLMs require vast amounts of data to function effectively, often including personal user information. The potential for data breaches or misuse of information is a significant risk, particularly in the mental health context, where breaches of confidentiality can have severe consequences for individuals. Ensuring robust data protection measures and clear, transparent communication about data usage is essential to building trust in LLM-based mental health services.

Trust and reliability issues

Building trust in LLMs as reliable sources of mental health support is another major challenge. For LLMs to be effective, users must believe that the information and advice they receive are accurate, safe, and beneficial. However, LLMs are not infallible; they can make mistakes, provide outdated information, or fail to understand complex human emotions and contexts. While AI can provide initial support, many users express hesitation about relying on it for more complex emotional or psychological issues (Gabarron, 2023). This skepticism is understandable, given that LLMs do not possess genuine empathy or understanding—they operate based on patterns in data rather than genuine emotional intelligence. Addressing these concerns involves improving the technical accuracy of LLMs and integrating them with human oversight to ensure that users receive the best possible care.

Ethical and cultural sensitivity

Ethical considerations are central to the deployment of LLMs in mental health. In a diverse country like India, where cultural norms and values vary widely, LLMs must be sensitive to these differences. Providing mental health support that does not align with cultural values can lead to mistrust or even harm to users. For example, advice that might be appropriate in a Western context could be inappropriate or even offensive in an Indian context. LLMs need to be trained on data that reflects the cultural nuances and ethical considerations relevant to the target population. This requires collaboration between AI developers, mental health professionals, and cultural experts to ensure the technology is effective and respectful of cultural differences (Binns, 2018).

OPPORTUNITIES FOR LARGE LANGUAGE MODELS IN WOMEN'S MENTAL HEALTH

Expanding access to mental health resources

One of the most promising opportunities LLMs present is the ability to expand access to mental health resources for women, particularly in underserved regions. LLMs can provide immediate, on-demand support, offering advice, information, and comfort to those without access to mental health services. This is especially important in rural areas of India, where mental health professionals are scarce, and the stigma surrounding mental health issues may prevent women from seeking help. Telepsychiatry and digital health interventions have significantly improved access to mental health care in remote areas (Haynes et al., 2024). LLMs could build on this success by providing scalable, low-cost support accessible to women nationwide. Integrating LLMs with existing telehealth platforms makes it possible to reach a larger population and provide more comprehensive care.

Personalized mental health support

LLMs have the potential to offer highly personalized mental health support tailored to the specific needs of each user. By analyzing user interactions, LLMs can adapt their responses to better suit individual preferences, cultural backgrounds, and emotional states. This level of personalization can make mental health support more effective, as it allows for more nuanced and relevant advice. For instance, an LLM could be trained to recognize signs of distress specific to Indian women, such as stress related to family responsibilities or societal pressures. LLMs can offer more meaningful support by providing advice specifically tailored to these challenges. Personalized digital interventions have been found to be more effective in improving mental health outcomes than generic advice (Hornstein et al., 2023).

Anonymity and confidentiality

One of the significant advantages of LLMs in mental health is the ability to provide support anonymously. Many women may feel uncomfortable discussing their mental health issues with family members or even with healthcare professionals due to stigma or fear of

judgment. LLMs can offer a confidential space where women seek help without revealing their identities. This anonymity can encourage more women to seek help, particularly in conservative communities where mental health issues are often stigmatized. Anonymity has been identified as a critical factor in individuals' willingness to seek online mental health support. By offering a safe, confidential space, LLMs can play a crucial role in breaking down the barriers to mental health care (Borghouts et al., 2021).

Scalability and cost-effectiveness

LLMs offer a scalable solution to the growing demand for mental health services. With a limited number of mental health professionals in India, especially in rural areas, LLMs can help bridge the gap by providing support to a large number of users simultaneously. This scalability is essential in public health initiatives, where reaching large populations with limited resources is often challenging.

Moreover, LLMs can be a cost-effective solution for providing mental health support. Once developed, these models can operate with minimal ongoing costs, making them accessible to a broader population segment. This cost-effectiveness is crucial in a country like India, where financial constraints often limit access to healthcare services.

Enhancing public health campaigns

LLMs can also play a significant role in enhancing public health campaigns focused on mental health. Integrating LLMs into existing public health infrastructure makes it possible to provide targeted mental health advice and support as part of broader health initiatives. For example, LLMs could be used to provide information and support during public health emergencies, such as the COVID-19 pandemic.

During the pandemic, digital health tools became essential in managing mental health crises. LLMs could be further developed to provide real-time advice and support during such crises, helping to reduce the mental health burden on the population. The potential of digital

interventions in public health campaigns has been recognized for their ability to reach large populations quickly and efficiently (Mason et al., 2022).

Culturally relevant interventions

Finally, LLMs offer the opportunity to develop culturally relevant mental health interventions. By training LLMs on data that reflects the cultural and societal context of Indian women, it is possible to create tools more in tune with local realities. This cultural relevance can make the support provided by LLMs more compelling and relatable, leading to better mental health outcomes. For example, an LLM could be trained to understand the specific cultural challenges faced by Indian women, such as the pressures of arranged marriages or the stigma associated with mental health issues. By providing advice sensitive to these cultural factors, LLMs can offer more meaningful and practical support.

Integrating culturally relevant LLMs in mental health care could also help address the unique challenges faced by women from different cultural backgrounds within India. This approach would ensure that the mental health support provided is not only practical but also respectful of the diverse cultural landscape of the country.

With their potential for scalability, cost-effectiveness, and personalization, LLMs offer a promising tool for enhancing mental health support for women in India. However, their successful implementation requires careful consideration of the challenges, including bias, accessibility, privacy, trust, and cultural sensitivity. By addressing these challenges, LLMs can become a valuable resource in improving mental health outcomes for women, particularly in underserved and marginalized communities.

Conflict of interest: The authors declare no conflict of interest.

Ethical considerations: Ethical approval is not applicable for this article.

Funding sources: None.

Authors' contributions: Both authors contributed equally to the conceptualization, resources, analysis, writing, and review of this manuscript.

References

- Balcombe, L., & De Leo, D. (2021). Digital mental health challenges and the horizon ahead for solutions. *JMIR Mental Health*, 8(3), e26811.
- Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29, 4349–4357.
- Borghouts, J., Eike, E., Mark, G., De Leon, C., Schueller, S. M., Schneider, M., ... & [others]. (2021). Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review. *Journal of Medical Internet Research*, 23(3), e24387.
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability, and Transparency*, 149–159.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91.
- Gabarron, E., Reichenpfader, D., & Denecke, K. (2023). Exploring the evolution of social media in mental health interventions: A mapping review. *Yearbook of Medical Informatics*, 32(1), 152–157.
- Haynes, S. C., Marcin, J. P., Yellowlees, P., Yang, S., & Hoch, J. S. (2024). Tele-psychiatry to address severe persistent mental illness in rural communities: An economic and break-even analysis. *Journal of Mental Health and Clinical Psychology*, 8(2), 35–41.
- Hornstein, S., Zantvoort, K., Lueken, U., Funk, B., & Hilbert, K. (2023). Personalization strategies in digital mental health interventions: A systematic review and conceptual framework for depressive symptoms. *Frontiers in Digital Health*, 5, 1170002.
- Mason, C., Lazenby, S., Stuhldreher, R., Kimball, M., & Bartlein, R. (2022). Lessons learned from implementing digital health tools to address COVID-19 in LMICs. *Frontiers in Public Health*, 10, 859941.
- Rana, U., & Singh, R. (2023a). Artificial intelligence for women's psychological wellbeing. *Psychiatria Danubina*, 35(3), 460–461.
- Rana, U., & Singh, R. (2023b, September 21). *The role of artificial intelligence in mental health care*. <https://doi.org/10.31235/osf.io/r4umy>

Correspondence:

Dr. Rupender Singh, Khalifa University,
Abu Dhabi, UAE
rupender.singh@ku.ac.ae

Published under



<https://creativecommons.org/licenses/by-nc-nd/4.0/>