Deploying Large Language Model-based Chatbots for Indic Languages in Low-Resource Regions

Anonymous submission

Abstract

There are many public health challenges in the so-called global south that are tackled through literacy-focused interventions by nonprofits and public health agencies. However, it is extremely challenging to scale such interventions to large populations in countries that speak multiple languages, and have complex sociocultural norms. We leverage large language models (LLMs) to demonstrate a scaling of successful literacy interventions to improve menstrual health and hygiene among adolescent females in Bangladesh. Our system is deployed as a WhatsApp chatbot that generates responses grounded in accurate, verified knowledge from international health agencies and domain-specific knowledge from local health organizations. We discuss the considerations for the development and deployment of consumer-facing LLM chatbots to be accessible in low-resource communities without a large English-speaking population.

Introduction

Many adolescent girls in rural areas of Bangladesh who have reached the age of menarche are unaware of the term 'menstruation' (Alam et al. 2022). This has influenced their outlook towards menstruation created a sense of inferiority amongst girls and also affected their school performance. Lack of accurate information has led to improper hygienic habits such as not knowing the correct methods to dispose of a sanitary pad. Traditional beliefs have cast menstruation in a negative light by considering it as something unhygienic and impure (Bhartiya and Aru 2013). This has changed the way society looks at this natural biological process.

Many initiatives have been taken to educate school-going girls about menstrual health and hygiene (Ambastha and Chowdhary 2023). Menstrual health interventions are coordinated activities, programs, or initiatives that are aimed at addressing the knowledge gap, stigma, and taboo that surrounds menstrual health. Such interventions aim to educate young girls and women about menstruation, and proper hygiene practices and make them self-aware about managing the menstrual cycle. These interventions help adolescent girls fight the anxiety associated with menstruation. It helps in regaining their self-esteem and confidence thereby leading to increased attendance in school and participation in other activities. A survey conducted in Bangladesh, suggests that menstrual literacy interventions have decreased the level of anxiousness in schools about menstruation(Alam et al. 2022). According to another survey performed by Haque et al. (Haque et al. 2014), there was a significant improvement in knowledge about menstruation starting from 51 percent baseline to 82.4 percent closing.

Menstrual Health and Hygiene in Bangladesh

Menstrual Health Management (MHM) is the process of dealing with menstruation with proper hygiene products with safety and dignity. Every day, more than 300 million women are going through their menstrual cycle (Bharath 2023). Many women and girls do not have the right information due to constrained education and the stigma around this subject. Only 6 percent of girls in Bangladesh have an MHM education and 53 percent of adolescent girls know about menstruation before their first period. Around 41 percent of Bangladeshi girls take an absence from their schools during their period (Star 2023). The WASH program released by BRAC (Bangladesh Rural Advancement Committee) brought a substantial improvement in creating facilities for women and girls in Bangladesh. This program led to a significant decrease in the absence rate of girls during their periods in intervention schools by 1 day compared to the girls in control schools and increased the likelihood of using hygiene products by 32 percent more than control school girls (Biswas et al. 2014). However, these solutions are labor intensive and very expensive to implement and maintain (Zillur et al. 2018-19). The information available on the Internet is often in English and does not provide solutions in Indic languages like Bengali (r21 2023). Therefore, leveraging generative Artificial Intelligence (AI), we develop a WhatsApp-based chatbot 'Sakhi'[1], which can communicate accurate MHM information in Bangla to cater to the 44 million WhatsApp users in Bangladesh (WorldPopulationReview 2023).

Scaling Literacy-focused Interventions

Chatbots powered by large language models present a means to provide interactive education on topics of interest in a direct manner to consumers of such information. These AI systems can hold natural conversations with users, understand their questions, and provide helpful information tai-

[1]Sakhi refers to 'female friend' in Hindi and Bangla

lored to their needs. Unlike static content or rule-based systems, chatbots allow users to ask follow-up questions and clarify points they don't understand at first. We designed our system with safety measures such as rate limiting, content filtering, and data privacy protections. By curating verified content from trustworthy sources, our system enables users to learn interactively at their own pace about sensitive topics like health and hygiene, while minimizing risks from misinformation. It is difficult to reduce these risks to zero, which is why we are intentional in our rollout of such a system to smaller groups for quality assurance, before making our system generally available to target communities that our NGO partner serves.

Risks of Large Language Models (LLMs)

While LLMs offer the ability to generate contextually relevant responses, concerns loom over potential pitfalls. One major apprehension is hallucination, where models may produce inaccurate or misleading information. Specifically, LLMs may generate confident assertions that lack grounding in verifiable evidence or are incompatible with established facts. This risks propagating misinformation if applied uncritically (Huang et al. 2023a). Factual errors in LLMs can arise when the model outputs information conflicting with the truth, and reasoning errors occur when it fails to provide logically consistent answers. (de Wynter et al. 2023) The costs associated with LLMs are substantial, with estimates ranging from hundreds of millions to billions of dollars for training, along with significant energy consumption and carbon emissions. Safety concerns include the potential for malicious use, lack of transparency in development, biases from training data, and the challenge of aligning LLMs' values with human values. (Tokayev 2023) Data privacy issues arise from the use of large volumes of web data, raising concerns about consent and anonymization. Additionally, the abuse potential of LLMs includes backdoor and prompt injection attacks, enabling manipulation of model behavior and the generation of harmful content, fake academic works, and malware. (Huang et al. 2023b)

Retrieval-augmented Generation

Retrieval-augmented generation (RAG) is an AI framework pioneered by former Meta researchers for improving the quality of LLM-generated responses by grounding the model on external sources of knowledge to supplement the LLM's internal representation of information (Martineau 2023). RAG aims to predict the output y based on the source input x (x, y are from a corpus D), while a document reference set Z is accessible (e.g., Wikipedia) (Yu 2022). RAG helps LLMs to generate more knowledgeable, diverse, and relevant responses by allowing them to access external data sources at inference time to build a richer prompt that includes context, history, and recent/relevant knowledge. RAG can significantly improve the accuracy of LLMs' responses, especially in cases involving domain-specific knowledge (Chen et al. 2023). It is a promising approach for mitigating the hallucination of LLMs, and it can outperform LLMs without retrieval by a large margin with much fewer parameters. (Manny 2023)

Deploying LLM Interventions

WhatsApp is a social media platform that allows users to exchange text, voice, and video messages. With over 2 billion users, it is one of the leading messaging services in the world (Dean 2023). Communication on WhatsApp is secured through end-to-end encryption. Bangladesh has close to 44 million WhatsApp users (WorldPopulationReview 2023). In addition to its ease of use, it has ubiquitous accessibility and widespread usage spanning socioeconomic divides. These requirements aligned with our target audience which were adolescent girls in rural areas of Bangladesh. From a product standpoint, WhatsApp has a service called WhatsApp Business. This service provides an API that helps in creating user-friendly template messages, automating responses, and listening for notifications using webhooks. It is easy to scale and develop conversational chatbots for WhatsApp, despite the manual approvals-related blockers involved in this process (for good reason, one might be inclined to add). Considering these tradeoffs, WhatsApp seemed to be the best service for deploying the bot.

Reproducibility

When deploying AI systems, it is critical to ensure reproducibility and extension of the system is trivially possible without dependencies of any kind on the original developers. The terms of the agreement we entered with our nonprofit partner have us commit to developing an architecture independent of proprietary tools, thus avoiding a lock-in with us as an external dependency to continue their work. While it is important to release open-source tools as a long-term goal, a good open-source library would require a significant commitment of resources towards documentation, code coverage, unit tests, and maintenance which is challenging for smaller teams with severely limited resources (like ours). In our case, this can be supplanted by building shareable, open-access tools that serve bespoke use cases for smaller organizations. The saved time could be spent on providing a seamless integration of our open-access tools with the external partners' offerings to the community. As broader use cases emerge with our early-stage open-access tools, we can share our codebase with more organizations. and jointly commit resources to develop documentation and necessary infrastructure to release it as an open-source library. This model has its limitations in that it reduces the scope of the AI tools deployed since the codebase may not be publicly accessible, but it also results in an earlier timeline for releasing system demos to relevant orgs., and it promotes working with partners to build out use cases for the AI system before sharing it with the open-source community.[2]

[2]An API key is a unique identifier that authenticates a user while using an API. For the Sakhi bot, we used OpenAI's API key, which is paid, to leverage the GPT-4 model. We maintained a public GitHub repository to keep track of the code. However, we accidentally ended up committing the key to the repository by including a historical commit with cached keys. Within a few minutes, the API key was exhausted, incurring a loss of USD 350 with our API provider, who–fortunately for us–waived part of this cost

Safety Measures to Prevent Abuse

In the pursuit of deploying Large Language Model (LLM) interventions for menstrual health and hygiene, ensuring user safety and preventing potential abuse are paramount considerations. To safeguard the system from misuse, we have implemented stringent safety measures. Firstly, a rate-limiting mechanism has been instituted, allowing a maximum of 20 requests per user within a 24-hour time frame. This not only prevents excessive utilization but also promotes fair and equitable access to the chatbot resources in the limit of concurrent requests that the LLM inference engine supports. Additionally, user data collection is conducted with a strong commitment to anonymity, prioritizing privacy and confidentiality, storing all data in an encrypted database. The collection of user chats is anonymized, respecting individual privacy while still gathering insights for research and development. These measures underscore our dedication to fostering a secure and respectful interaction environment for users, aligning with ethical principles and data protection standards.

Introducing our Chatbot to the community

Along with collecting data from verified resources such as UNESCO, we also verify the legitimacy of the responses given by the chatbot with the help of certified healthcare professionals. Our chatbot, Sakhi, was built in partnership with a UK-based charity called WaterAid which has worked tirelessly in Bangladesh to create a difference. With the help of Wateraid, we can leverage trust in a huge community of people who can benefit from this service. Our privacy policy and terms and conditions are written in Bangla, the most spoken language in Bangladesh. Our privacy policy adheres to the rule of maintaining user anonymity by only collecting data without the user's Personally Identifiable Information (PII). Our Terms and Conditions state the authority over the user conversations and not their personal information.

Deployment

A robust deployment architecture was critical to enable real-time conversational interactions between users and the question-answering system. The deployment process involved setting up and integrating several components to facilitate the inference pipeline, expose it securely over the internet, and handle user interactions via a chatbot interface.

Configuring Flask App Inference Pipeline

The RAG model for question answering was incorporated into a Flask application that handled the inference pipeline execution. The Flask (https://flask.palletsprojects.com/en/3.0.x/) app orchestrated the overall flow receiving input questions, triggering the RAG model to generate responses by retrieving relevant information, and returning the responses. This server-side application handled the bulk of the logical processing.

at our request.

Ngrok Integration

To allow external access to the Flask application server running locally, the Ngrok (https://ngrok.com/) secure tunneling platform was leveraged. Ngrok exposed the Flask app to the internet by establishing a secure tunnel. This enabled other systems to access the inference pipeline over the web using the Ngrok public URL.

Google Cloud Hosting

The Flask application server handling the inference pipeline was hosted on the Google Cloud Platform (GCP) to enable scalable and reliable performance. GCP provided a managed hosting environment to run the Flask app, ensuring it could handle high query volumes from users without disruption.

Webhook Integration

A webhook was implemented to facilitate integration with the chatbot interface. The webhook created a public endpoint to receive incoming messages from the chatbot. It was configured to process these messages, trigger the inference pipeline on the Flask app, and return the responses.

WhatsApp Chatbot Configuration

The conversational interface for end-users was enabled through a WhatsApp chatbot. This chatbot was integrated with the aforementioned webhook to communicate in real-time with the inference pipeline. The WhatsApp platform was configured to forward any messages to the webhook endpoint, which would process the user question, generate a response, and send it back to the chatbot to display.

References

2023. Most used languages online by share of websites 2023. https://www.statista.com/statistics/262946/most-common-languages-on-the-internet/. Accessed: 2023-11-28.

Alam, M.-U.; Sultana; Farhana; Hunter; C, E.; Winch; J, P.; Unicomb; Leanne; Sarker, S.; Mahfuz, M. T.; Al-Masud, A.; Rahman, M.; and Luby, S. P. 2022. Evaluation of a menstrual hygiene intervention in urban and rural schools in Bangladesh: a pilot study. *BMC public health*, 22(1): 1–16.

Ambastha, N.; and Chowdhary, R. S. 2023. Government endeavor towards improving Menstrual Health and Hygiene in the State of Bihar.

Bangladesh, W. 2014. advancing menstrual hygiene. https://www.wateraid.org/bd/sites/g/files/jkxoof236/files/briefing-paper-on-menstrual-hygiene-lets-talk-about-it-.pdf. [Online; accessed May-2014].

Bharath, S. 2023. SEEN, BUT IGNORED: A CRITICAL REVIEW ON MENSTRUAL HYGIENE PRACTICES IN INDIA. *EPRA International Journal of Research and Development (IJRD)*, 8(8): 110–119.

Bhartiya; and Aru. 2013. Menstruation, religion and society. *International Journal of Social Science and Humanity*, 3(6): 523.

Biswas, D.; Shoab, A.; Mahbub-Ul Alam, L. U.; Sultana, M. M.; ur Rahman, M.; and Opel, A. 2014. Impact Assessment WaterAid of WASH in School Project. https://www.wateraid.org/bd/sites/g/files/jkxoof236/files/2020-07/WASH%20in%20School-MHM%20and%20Learning%20Impact%20-WAB.pdf. [Online; accessed July-2014].

Chen, J.; Lin, H.; Han, X.; and Sun, L. 2023. Benchmarking Large Language Models in Retrieval-Augmented Generation. arXiv:2309.01431.

de Wynter, A.; Wang, X.; Sokolov, A.; Gu, Q.; and Chen, S.-Q. 2023. An evaluation on large language model outputs: Discourse and memorization. *arXiv preprint arXiv:2304.08637*.

Dean, B. 2023. WhatsApp 2023 user statistics: How many people use WhatsApp? https://backlinko.com/whatsapp-users. Accessed: 2023-11-28.

Haque, S.; Rahman, M.; Itsuko, K.; Mutahara, M.; and Sakisaka, K. 2014. The effect of a school-based educational intervention on menstrual health: an intervention study among adolescent girls in Bangladesh. *BMJ Open*, 4(7). WOS:000339720900010.

Huang, L.; Yu, W.; Ma, W.; Zhong, W.; Feng, Z.; Wang, H.; Chen, Q.; Peng, W.; Feng, X.; Qin, B.; et al. 2023a. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *arXiv preprint arXiv:2311.05232*.

Huang, X.; Ruan, W.; Huang, W.; Jin, G.; Dong, Y.; Wu, C.; Bensalem, S.; Mu, R.; Qi, Y.; Zhao, X.; et al. 2023b. A Survey of Safety and Trustworthiness of Large Language Models through the Lens of Verification and Validation. *arXiv preprint arXiv:2305.11391*.

Jahan, H.; Ahmed, M.; and Roomy, S. 2014. Our journey towards making Menstruation a normal fact of life. https://www.wateraid.org/bd/sites/g/files/jkxoof236/files/2023-06/Bangladesh%20MHM%20Journey.pdf. [Online; accessed May-2014].

Manny, S. 2023. Retrieval augmented generation: Keeping LLMs relevant and current. https://stackoverflow.blog/2023/10/18/retrieval-augmented-generation-keeping-llms-relevant-and-current/. Accessed: 2023-11-23.

Martineau, K. 2023. What is retrieval-augmented generation? https://research.ibm.com/blog/retrieval-augmented-generation-RAG. Accessed: 2023-11-23.

Nuzhat, S.; Iyer, R.; Qader, A. I. A.; and Manzoor-Al-Islam, S. 2023. A Participatory Assessment for Climate-Induced WASH Vulnerabilities in Bangladesh. https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/18060/SLH_LB16_Participatory_Assessment_Climate-Induced_WASH_Vulnerabilities_Bangladesh.pdf?sequence=4&isAllowed=y. [Online; accessed July-2023].

Qader, A. I. A.; Abbas, F.; Manzoor-Al-Islam, S.; Shaikh, P. H.; and Jahan, H. 2021. WASH and climate: Policy and financing (dis)connects in Bangladesh. https://www.wateraid.org/bd/sites/g/files/jkxoof236/files/wash-and-climate---policy-and-financing-disconnects-in-bangladesh.pdf. [Online; accessed 31-March-2021].

Star, D. 2023. Period poverty: A silent crisis in Bangladesh. https://www.banginews.com/web-news?id=625036923120193102db9c3d4c4e5a1fc4c0edde. Accessed: 2023-11-23.

Tokayev, K.-J. 2023. Ethical Implications of Large Language Models A Multidimensional Exploration of Societal, Economic, and Technical Concerns. *International Journal of Social Analytics*, 8(9): 17–33.

WorldPopulationReview. 2023. WhatsApp users by country 2023. https://worldpopulationreview.com/country-rankings/whatsapp-users-by-country. Accessed: 2023-11-23.

Yu, W. 2022. Retrieval-augmented generation across heterogeneous knowledge. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop*, 52–58.

Zillur, H.; Mohammad; Rahman; Muhammad; Wazed, A.; Islam; Aminul; and Bangladesh, W. 2018-19. Bangladesh analysis on WASH budget allocation & expenditure FY 2018-19. https://www.wateraid.org/bd/sites/g/files/jkxoof236/files/2020-09/WASH%20Budget%20Research%20FY2018-19%20Full%20Report.pdf. Accessed: 2023-11-23.

Appendix

Deployment of AI tools in our case has been with an external organisation with whom we had no prior affiliation nor any funding agreements. We provide this section to discuss their history of work in an effort to spotlight the variety of areas that a deployed AI system like Sakhi can be adapted to support. For example, we discuss challenges related to sanitation and hygiene that are caused by a lack of awareness of accurate information and access to hygiene products, which could benefit from a similar scaling of digital literacy interventions as in the case of MHM discussed in the main paper. We present these areas to motivate the research community to invest in developing better AI tools for social good in low and middle-income countries.

Identifying Partners to Deploy AI: WaterAid

An international non-profit organization called WaterAid works to improve access to clean water, hygienic practices, and sanitation in some of the world's poorest communities to change lives. WaterAid was established in 1981 and is present in many nations, striving to guarantee that everyone has access to clean water, sanitary facilities, and appropriate hygiene habits.

How do they operate?

WaterAid's mission in Bangladesh is to improve access to safe water, sanitation, and hygiene (WASH) to transform lives. The organization employs multiple approaches to tackle the various obstacles that communities across the nation encounter. The organization works closely with local partners, government agencies, and community leaders

to develop plans that are tailored to the specific needs of the community, starting with thorough assessments of those needs. WaterAid prioritizes infrastructure development during project implementation, utilizing community-led strategies such as Community-Led Total Sanitation (CLTS) to guarantee engaged community involvement. Training courses and campaigns promoting hygiene awareness help create long-lasting behavioral changes. Governments are involved in advocacy efforts to create supportive policies, and a sense of ownership is fostered through community empowerment. Adaptive strategies are made possible by ongoing learning and innovation, including pilot projects. Thorough monitoring and assessment guarantee the effectiveness of initiatives, and collaborations with non-governmental organizations and international organizations enhance the overall effect, demonstrating WaterAid's dedication to enhancing WASH circumstances.

Evaluating Past Challenges and Successes to Inform AI Deployment

WaterAid's work in Bangladesh, like many similar organizations, encounters challenges such as limited resources, complex socio-cultural norms, and the need for sustainable solutions. These challenges are particularly relevant when addressing issues related to menstrual health and hygiene, where cultural taboos and lack of awareness pose significant obstacles. Presented below are some of the challenges that WaterAid has successfully navigated in Bangladesh:

WASH Hygiene Interventions: Many communities in Bangladesh face challenges in accessing safe and clean water sources, leading to waterborne diseases and poor health outcomes. WaterAid has implemented various WASH projects in Bangladesh to improve hygiene practices and access to clean water. These interventions often target communities with limited resources, addressing the root causes of poor sanitation and hygiene.

Community-Led Total Sanitation (CLTS): Inadequate sanitation facilities contribute to open defecation, leading to the spread of diseases and environmental pollution. WaterAid has been involved in promoting community-led approaches to sanitation. CLTS empowers communities to take collective action to become open-defecation-free, promoting the construction of latrines and changing hygiene behaviors.

School WASH Programs: According to the 2014 National Hygiene Baseline Survey, 57% of Bangladesh's secondary schools do not have individual restrooms for their students. Those who have private restrooms often deal with practical problems like broken toilets or unusable conditions brought on by improper upkeep and operation of the facilities et. al. (Biswas et al. 2014). Recognizing the importance of WASH in schools, WaterAid has worked on initiatives to provide clean water and sanitation facilities in educational institutions to more than 107,000 students. Improving WASH infrastructure in schools contributes to a healthier learning environment.

Hygiene Education and Behavior Change: Changing entrenched hygiene behaviors can be challenging, especially in areas with deeply rooted cultural practices. As per the MHM report, it was discovered that 80% of women and teenagers in their households used old clothes during menstruation. But women frequently don't know how important it is to wash these clothes in soapy water and let them air dry under the sun. WaterAid conducts hygiene education programs to raise awareness about the importance of good hygiene practices. This includes educating communities about handwashing, menstrual hygiene, and the overall link between sanitation and health. Due to such awareness sessions, girls from intervention schools had a 57% lower chance of missing school during menstruation, as compared to girls from control schools as per (Biswas et al. 2014)

Climate Resilience Initiatives: Bangladesh is vulnerable to climate change, leading to issues like increased salinity in water sources et. al. (Nuzhat et al. 2023). In Bangladesh, 70% of people reside in rural areas that are severely impacted by climate change. The budget allocation in this report by (Qader et al. 2021) favors urban areas over rural ones, making the latter more susceptible to the effects of salinization and climate change. WaterAid identifies such issues and adapts its projects to address climate-related challenges, such as developing resilient water sources that can withstand changing environmental conditions.

With a focus on urban areas, WaterAid has implemented projects to enhance water and sanitation services in crowded and underserved urban communities. These efforts aim to address the unique challenges of rapid urbanization. Through focused WASH projects, WaterAid addresses these issues and improves the general health and well-being of communities in Bangladesh. This paves the way for more comprehensive public health initiatives, like promoting menstrual health and hygiene et. al. (Bangladesh 2014)

Identifying use-cases to deploy AI

Future Goals and how they can be scaled cost-effectively? The evolution of Menstrual Health Management (MHM) initiatives, including WaterAid's approach, reflects a progressive shift from de-stigmatizing menstruation et. al. (Jahan, Ahmed, and Roomy 2014) to leveraging technology, particularly chatbots, for widespread impact. Advocacy efforts and community engagement have worked to normalize menstruation, challenging cultural taboos. With the increasing accessibility of technology, MHM initiatives have embraced chatbots as transformative tools, providing discreet and personalized support through widely used platforms. Recognizing the potential, WaterAid aims to integrate chatbots to cut costs, improve efficiency, and reach a larger population promptly. This aligns with the organization's commitment to innovative, sustainable solutions that empower communities, especially among the younger population accustomed to digital communication. Chatbots offer a cost-effective means of scaling interventions, streamlining communication, ensuring privacy, and overcoming geographical barriers, thus contributing to the broader goal of enhancing menstrual health and well-being in diverse communities.