LLM-enhanced Health Monitoring and Intervention for People with Dementia: Data Extraction and Conversation Design D.S. Kim, Y. Won, and I. Jung. *Gerontechnology 25*(s)

Purpose With nearly 50% of older adults diagnosed with dementia living alone in Korea, there is a growing need for communication agents such as AI speakers and robots to support health monitoring and self-help management. To utilize language as a new vital sign, systems must be able to detect health risk signals from everyday conversations, enabling suitable intervention; connections to medical services, health education, and timely interventions. **Method** Using a conversational dataset collected from 108 older adults with dementia, this study aims to extract health-related issues and design interventional dialogues. The dataset, compiled by the first author in the previous research, consists of 6,400 episodes and 13,000 transcribed utterances from conversations. In the preprocessed conversational data, health-related terms are first tagged using a Named Entity Recognition (NER) approach. N-grams consisting of five words before and after each tagged term are combined with semantic similarity analysis using BERT embeddings to establish rules for detecting potential risk signals. A set of Q&A dialogues based on the conversational dataset is used to conduct instruction fine-tuning on small Large Language Models (LLM). After comparing the performance of multiple small LLMs, the model with the lowest perplexity score is selected as the base model. These dialogues are designed to simulate conversations between an AI agent specialized in geriatrics and nursing and older adults at risk of dementia. Fine-tuning is performed iteratively until the model achieves the desired balance: producing domain-specific, professional-quality responses while preserving its ability to engage in natural, general-purpose dialogue. Once the chatbot is fine-tuned, it is deployed on a cloud server for real-world testing. **Results** This study developed a logic for identifying health risk signals by combining a rule-based approach—utilizing keywords, grammar, and context—with a large language model (LLM). First, key candidate sentences with high recall were detected using core keywords such as ‘hospital’, ‘sleep’, ‘urine’, ‘fall’, and ‘pain’. The study also identified several health risk categories in the dataset, including chronic diseases, sleep disorders, fall risk, nutritional issues, pain management and emotional well-being. Corresponding intervention answers were generated, covering medical facility referrals, recommendations for hydration and nutritional management, exercise and activity suggestions, and medication management checks. In this study, a chatbot was developed by integrating a categorized framework of common health issues in older adults with a domain-specific knowledge base in geriatrics and nursing. This integration enables the chatbot to detect potential health risks and provide appropriate, context-aware responses tailored to the needs of older users. The performance of the chatbot is initially evaluated by comparing chatbot’s responses and baseline GPT models using ROUGE-1, 2, and ROUGE-L scores. Higher similarity scores are interpreted as the chatbot more accurately reflecting the intended domain-specific expertise. **Discussion** Conversations with AI agents hold significant potential for health monitoring and intervention for people with dementia. However, current interactions with these agents often lack sufficient engagement, limiting the ability to extract meaningful health-related information. This study proposes a novel method to enhance engagement by training conversations using pre-existing dialogue data and developing an algorithm based on a hybrid approach that combines rule-based methods and LLM. The usefulness and effectiveness of the proposed conversational intervention program require further evaluation. Besides, the use of generative health risk interventions needs to address ethical issues, including safety and responsibility. A future task involves addressing these issues, as well as enhancing conversational immersion and examining its impact on self-management behaviors.

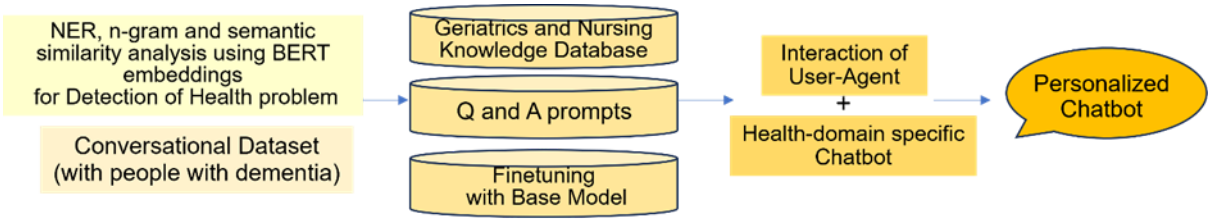

Figure 1. Workflow diagram

References

1. Ministry of Health and Welfare. (2023). 2023 National Survey on Dementia Epidemiology and Current Status in Korea. Sejong: Ministry of Health and Welfare.
2. Nouman, M., Sara, H., Khoo, S. Y., Mahmud, M. P., & Kouzani, A. Z. (2023, June). Mental health prediction through text chat conversations. In 2023 International Joint Conference on Neural Networks (IJCNN) (pp. 1-6). IEEE.

Keywords: Large Language Model (LLM), Health Monitoring and Intervention, Conversation Agents, People with Dementia, Health Conversation Design

Address: Department of Culture and Technology, Hankook University of Foreign Studies, S. Korea
Email: weeny388@gmail.com

Acknowledgement: This survey was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea [NRF-2024S1A16022129 to DS.K]