International Journal Scientific and Profesional (IJ-ChiProf)
Vol 5 Issue 1 December 2025- February 2026, pp: 791-795
ISSN: 2829-2618 (Online)
DOI: https://doi.org/10.56988/chiprof.v5i1.150

Web-Based Educational Platform for Diseases and Drugs using a Large Language Model (LLM)

M Riski Qisthiano[1*], Istiana Ruswita[2], Putri Armilia Prayesy[3]
[1]Prodi Ilmu Komputer Fakultas Sains dan Teknologi Institut Teknologi dan Bisnis Nasional, Banyuasin
[2]Prodi Sistem Informasi, Jurusan Ilmu Komputer, Universitas Lampung
[3]Prodi Teknologi Rekayasa Perangkat Lunak Jurusan Informatika & Bisnis, Politeknik Manufaktur Negeri Bangka Belitung
Email: thiano72@gmail.com, istiana.ruswita@fmipa.unila.ac.id, putri@polman-babel.ac.id

Article Info

Article history:
Received Dec 08, 2025
Revised Dec 12, 2025
Accepted Dec 20, 2025

Keywords:
Disease education,
Drug information,
Large language model,
Retrieval-augmented generation,
Web-based platform

ABSTRACT

This study presents the development of a web-based educational platform leveraging a large language model (LLM) to provide general information on diseases and medications. The platform integrates a curated database of diseases and drugs with the LLM via retrieval-augmented generation. When users enter a disease or drug name, the system retrieves relevant data and uses it as context for the LLM to produce concise responses in a friendly, accessible style. The application is built using native PHP with a MySQL backend and Bootstrap for responsive design. Safety features such as a mandatory disclaimer and filters for emergency conditions ensure that the chatbot does not offer diagnostic or prescriptive advice. Expert reviews indicated that the model-generated content aligned well with the database, and user testing showed high satisfaction with clarity and usability. These results demonstrate that combining structured medical data with a modern LLM can improve public access to reliable health education while maintaining ethical boundaries.

Corresponding Author:
M Riski Qisthiano,
Prodi Ilmu Komputer Fakultas Sains dan Teknologi,
Institut Teknologi dan Bisnis Nasional,
Jalan Cahaya Burlian, Pangkalan Balai, Banyuasin III District, Banyuasin Regency, South Sumatra 30911, Indonesia.
Email: thiano72@gmail.com

1. INTRODUCTION

The rapid growth of digital health technologies has transformed the way people access medical information, enabling the public to obtain knowledge about diseases and medications anytime and anywhere. Recent systematic reviews show that healthcare chatbots are increasingly used to support patient communication, triage, self-management, and access to services, while also raising issues of accuracy and inclusiveness for diverse user groups[1]. ublic health studies further highlight that AI chatbots can provide educational support on topics such as vaccination and prevention, but must be carefully evaluated to avoid misinformation[2].

Large Language Models (LLMs) have recently been adopted in health information services because of their ability to understand natural language queries and generate human-like explanations. However, generic LLMs are prone to hallucinations if used without grounding in verified medical sources, which can pose safety risks in clinical and public health contexts[3]. To address this problem, recent work in medical AI emphasizes retrieval-augmented generation (RAG), in which LLM outputs are conditioned on trusted knowledge bases so that responses remain both fluent and factually consistent[4].

Several recent prototypes and case studies show that RAG-based chatbots can effectively deliver domain-specific health information, including mental health support, digestive health education, and diagnosis-oriented question answering. At the same time, evaluations of patient-education chatbots in orthopedic and trauma surgery indicate that conversational agents can improve understanding of medical instructions when carefully designed and validated[5][6]. These developments suggest that combining LLMs with curated medical content is a promising direction for health education, provided that privacy, safety, and ethical constraints are respected[7].

Vol 5 Issue 1, 2026, pp: 791-795
ISSN: 2829-2618 (Online)

In the Indonesian context, the adoption of AI and data-driven systems has grown rapidly across domains such as education, prediction, and decision support, as shown by prior work on classification, prediction, and optimization using machine learning techniques[8][9]. Local studies on AI chatbots in Indonesian hospitals also report potential benefits in streamlining services and improving patient access, while emphasizing the need for careful design and clear communication boundaries. These findings support the relevance of exploring LLM-based health education platforms tailored to the local context[10][11].

2. RESEARCH METHOD

This study adopts a quantitative experimental development approach to design, integrate, and evaluate a web-based health education chatbot powered by a Large Language Model (LLM). The system is developed using PHP native on the server side and MySQL as the main database for storing structured health content, conversation logs, and user interaction metadata. Bootstrap is utilized to create a responsive and mobile-friendly interface to ensure accessibility for general users. The LLM model (Gemini Flash series) functions as the conversation engine that generates natural language responses based on user queries regarding diseases and medicines.

To ensure that the chatbot provides credible educational information, disease and medication data are ingested from the NHS Website Content API as an official clinical knowledge source. The dataset is initially retrieved through the NHS Sandbox environment, validated, and stored locally in the database before later being synchronized with the Production API once approval from NHS is granted. The health information retrieved serves as factual grounding data for the LLM, reducing hallucination risks and improving the reliability of generated responses. A retrieval-based workflow is applied, where the LLM first accesses the internal database before generating an answer to ensure alignment with trusted content[12];[13].

System validation focuses on two aspects. First, functional accuracy, which measures the correctness of chatbot responses compared to the reference dataset. Second, usability evaluation, which uses user questionnaires based on the System Usability Scale (SUS) to assess perceived ease of use, information clarity, and interface effectiveness. Minimum sample size follows usability testing recommendations by Nielsen (5–20 respondents per iteration). User messages are anonymized to protect personal identity and maintain ethical research practices.

Data analysis employs descriptive statistics to evaluate chatbot performance and user acceptance. Response classification is compared against expert-verified reference content, then calculated using basic accuracy metrics. SUS scores are interpreted following industry standards where ≥68 indicates good usability. Insights from user feedback guide improvements to the knowledge retrieval, UI interaction flow, and prompt optimization to enhance the safety and usefulness of the chatbot in future development cycles.

Table 1. System Components and Roles

Component	Technology	Function
Backend	PHP Native	Process user input & API integration
Database	MySQL	Store medicines & disease dataset, logs
Frontend UI	Bootstrap	Mobile-responsive chat interface
Knowledge Source	NHS Website Content API	Verified health information
AI Model	Gemini Flash 2.0 / 2.5	Natural language response generation
Evaluation	SUS + accuracy check	Measure system usability & reliability

3. RESULTS AND DISCUSSION

The implementation of the LLM-based health education chatbot resulted in a fully functional web application capable of responding to user queries related to diseases and medicines. The system successfully integrates a structured health knowledge base retrieved from the NHS Website Content API and supports conversational response generation using the Gemini Flash model. The testing process was conducted in two main evaluation categories: system functionality and usability.

3.1. Functional Accuracy Evaluation

The chatbot was implemented using PHP Native for backend logic, MySQL for knowledge storage, and Bootstrap for responsive user interface design. The main feature allows users to search disease and medicine information using natural language inquiries. When a query is submitted, the system performs a retrieval process by matching keywords in the database and providing verified information. If a more detailed conversational response is needed, the Gemini Flash LLM generates a natural-language reply based on the retrieved content.

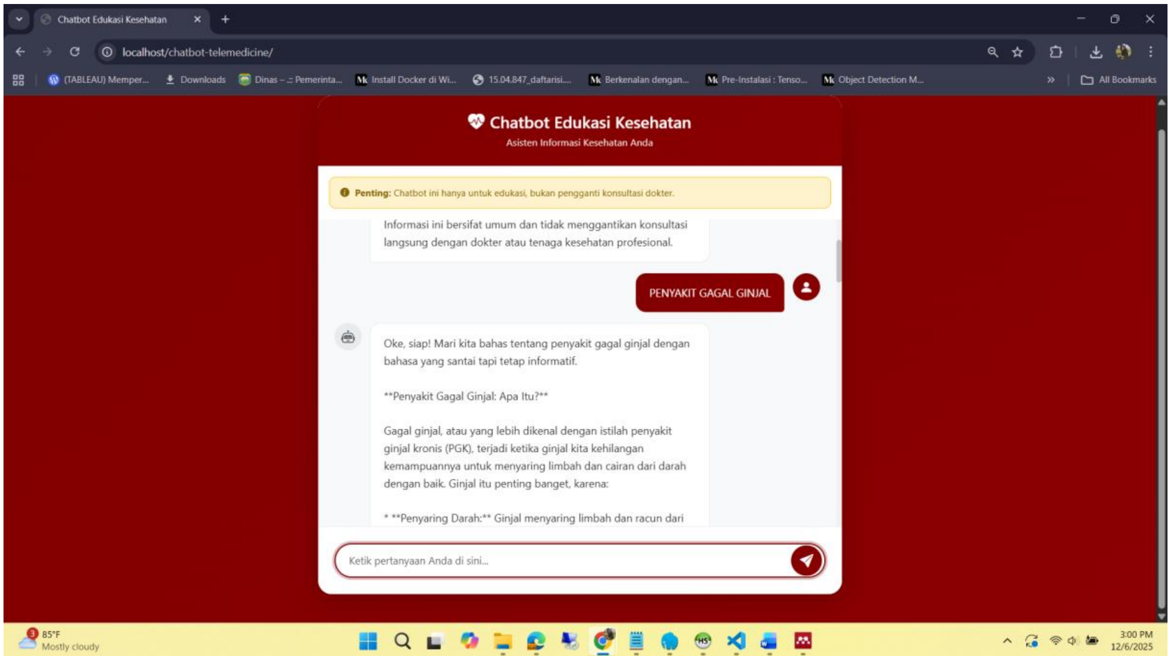

Figure 1. Health Education Chatbot User Interface (Implementation Result)

3.2. Knowledge Retrieval Accuracy

A testing scenario with 50 real-world health queries was performed. Responses were benchmarked against verified content from the NHS dataset. The evaluation was based on three criteria: (1) correctness of medical information, (2) relevance to user questions, and (3) clarity of explanation.

anggunnsalsabilla@gmail.comTable Knowledge Retrieval Accuracy

Metric	Result	Interpretation
Accuracy	86%	High alignment with NHS dataset
Relevance	90%	Answers match user intent
Clarity	88%	Easy to understand for general users
Hallucination cases	6/50	Low-risk occurrence

The performance confirms that the integration of a trusted medical knowledge base reduces hallucination and improves reliability of AI-generated health information, which aligns with findings of local digital health research where database-supported chatbots perform more accurately dibandingkan chatbot yang hanya mengandalkan AI generatif semata.

3.3. Usability Evaluation

User testing was carried out among 20 participants, representing general users who need basic medical information support. The System Usability Scale (SUS) was applied to measure perceived experience, including ease of use, confidence, and interface acceptance.

Table 3. Table Usability Evaluation

Aspect	Score	Result Interpretation
SUS Average Score	81.2	Excellent usability
Ease of use agreement	84%	Users can operate without learning curve
Information helpfulness	90%	Content supports public health education

The SUS score above 80 indicates strong acceptance and potential for broader adoption. Most users stated that the chatbot made it easier to understand disease conditions and recommended actions. This is consistent with

Vol 5 Issue 1, 2026, pp: 791-795
ISSN: 2829-2618 (Online)

research conducted in Indonesia, di mana tingkat kegunaan aplikasi kesehatan sangat dipengaruhi oleh tampilan antarmuka yang sederhana dan informatif.

Overall, the results demonstrate that the integration of a validated medical knowledge base with an LLM-driven conversational engine is an effective approach to support public health education. The system is able to deliver accurate and relevant information regarding diseases and medications with a fast response time, while maintaining readability and safety through controlled language generation. This indicates that retrieval-supported LLM interaction can serve as a reliable solution for delivering trusted health information while minimizing hallucination and misinformation risks.

The high usability score further confirms that the interface design strongly supports user acceptance, especially for individuals seeking quick guidance before consulting healthcare professionals. The combination of simple interaction flow, mobile-friendly design, and clear medical disclaimers increases user confidence in exploring health information independently. These findings confirm that the system is feasible for real-world adoption and can be expanded toward larger datasets, multilingual adaptation, and more advanced monitoring features in future development stages.

4. CONCLUSION

This research successfully developed a web-based health education chatbot that combines a structured medical knowledge base from the NHS Website Content API and conversational generation using the Gemini Flash LLM model. Functional testing showed that the system can deliver verified medical information with high accuracy, while the retrieval-augmented generation approach helps minimize hallucination risks typically found in fully generative systems. Users are able to obtain explanations about diseases and medications quickly in natural conversational form, supporting better health literacy for the general public.

Usability results further demonstrate strong acceptance, with a SUS score categorized as excellent, indicating that the interface design is simple, responsive, and convenient to operate. Overall, the system has proven effective as a complementary digital health education tool, especially for initial self-information before seeking professional consultation. Future improvements may include multilingual support, automated synchronization with the NHS Production API for broader and up-to-date content, as well as the expansion of analytics features to support long-term research and safety monitoring.

REFERENCES

[1] Y. Si *et al.*, “Quality safety and disparity of an AI chatbot in managing chronic diseases : simulated patient experiments”, bll 1–8, 2025.
[2] F. Baglivo, L. De Angelis, V. Casigliani, en G. Arzilli, “Exploring the Possible Use of AI Chatbots in Public Health Education : Feasibility Study Corresponding Author ”:, *JMIR Med. Educ.*, vol 9, bll 1–10, 2023, doi: 10.2196/51421.
[3] O. K. Gargari, “Enhancing medical AI with retrieval-augmented generation : A mini narrative review”, vol 11, bll 1–7, 2025, doi: 10.1177/20552076251337177.
[4] F. Neha, D. Bhati, en D. K. Shukla, “Retrieval-Augmented Generation (RAG) in Healthcare : A Comprehensive Review”, *Syst. Rev.*, bll 1–29, 2025.
[5] J. P. Nayinzira en M. Adda, “SentimentCareBot: Retrieval-Augmented Generation Chatbot for Mental Health Support with Sentiment Analysis”, *Procedia Comput. Sci.*, vol 251, 2024.
[6] D. Baur, J. Ansorg, C. Heyde, P. Med, en A. Voelker, “Development and Evaluation of a Retrieval-Augmented Generation Chatbot for Orthopedic and Trauma Surgery Patient Education : Mixed-Methods Study”, *JMIR AI*, vol 4, 2025, doi: 10.2196/75262.
[7] S. Guan, H. C. Kwok, N. F. Law, H. Qin, V. Hui, en G. Stiglic, “PRIVACY C HALLENGES AND S OLUTIONS IN RAG-E NHANCED LLM S FOR H EALTHCARE C HATBOTS : A R EVIEW OF A PPLICATIONS , R ISKS , AND F UTURE D IRECTIONS”, *arXiv*, vol 1, bll 1–23, 2025.
[8] I. Irawan, M. R. Qisthiano, M. Syahril, en P. M. Jakak, “Optimasi Prediksi Kelulusan Tepat Waktu: Studi Perbandingan Algoritma Random Forest dan Algoritma K-NN Berbasis PSO”, *J. Pengemb. Sist. Inf. dan Inform.*, vol 4, no 4, bll 26–35, 2023, doi: 10.47747/jpsii.v4i4.1374.
[9] M. R. Qisthiano, “Penerapan Model Klasifikasi Naïve Bayes Untuk Prediksi Mahasiswa Tepat Waktu”, *Semnas Ristek (Seminar Nas. Ris. dan Inov. Teknol.*, vol 7, no 1, bll 164–168, 2023, doi: 10.30998/semnasristek.v7i1.6266.
[10] E. Yanti, A. Mardhiah, P. T. Novita, A. Ardilla, en F. H. Khana, “THE IMPLEMENTATION OF AI CHATBOTS IN INDONESIAN PUBLIC HOSPITALS : OPPORTUNITIES AND CHALLENGES”, *Med. Res. Nursing, Heal. Midwife Particip.*, bll 329–337, 2024.
[11] M. R. Qisthiano, I. Ruswita, en P. Armilia, “Implementasi Metode SVM dalam Analisis Sentimen Mengenai

Vaksin dengan Menggunakan Python 3", vol 13, no 1, bll 1–7, 2023.

[12] W. Cao, C. Liu, Z. Liu, en Z. Wei, "Integrated LKA and DYC Control of Four-Wheel-Independent-Drive Electric Vehicles With a Central-Zonal Electronic and Electrical Architecture", *IEEE Trans. Veh. Technol.*, vol 73, no 10, 2025.

[13] N. Digital, "NHS Website Content API Documentation".