KOMPUTASI: JURNAL ILMIAH ILMU KOMPUTER DAN MATEMATIKA
VOL. 23 (2) (2026), XX-XXp-ISSN: 693-7554, e-ISSN:2654-3990

Implementation of Artificial Intelligence in Health Screening Systems for Category-Based Fitness and Nutrition Recommendations

Nasrul[1], Henry Saptono[2], Rusmanto[3]

[1,2] Department of Informatics Engineering, Nurul Fikri Integrated School of Technology (Sekolah Tinggi Teknologi Terpadu Nurul Fikri), West Java, 16451, Indonesia

[3] Department of Information Systems, Nurul Fikri Integrated School of Technology (Sekolah Tinggi Teknologi Terpadu Nurul Fikri), West Java, 16451, Indonesia

Abstract

The increasing prevalence of non-communicable diseases, such as diabetes, hypertension, and metabolic disorders, highlights the need for accessible health screening services. However, existing health screening systems generally provide examination results without offering automated educational recommendations to support preventive healthcare. This study proposes an Artificial Intelligence (AI)-based health screening information system that automatically generates fitness and nutrition recommendations based on categorized health screening results. The research employed a Research and Development (R&D) approach using the Extreme Programming (XP) software development methodology. The proposed system was developed using the Laravel framework and the Filament administration panel and integrates a Large Language Model (LLM) through the OpenAI API. The proposed architecture combines structured prompt engineering, predefined AI guardrails, SHA-256 prompt hashing, and recommendation caching to improve recommendation consistency and computational efficiency. The system was evaluated using User Acceptance Testing (UAT) and an AI Recommendation Consistency Evaluation. The UAT results showed that all functional requirements were successfully fulfilled. The consistency evaluation demonstrated that repeated processing of identical health screening data produced stable recommendations, achieving an average qualitative consistency score of 90.4% and an average TF–IDF cosine similarity of 0.784. These findings indicate that the proposed architecture is capable of generating reliable and consistent educational recommendations while reducing redundant AI requests through recommendation caching. This study contributes to the development of intelligent health information systems by introducing an efficient AI recommendation architecture that supports digital health screening and preventive healthcare services.

Keywords: *Artificial Intelligence; Health Screening Information System; Prompt Engineering; Recommendation Caching; Preventive Healthcare.*

1. Introduction

Non-communicable diseases (NCDs), including diabetes mellitus, hypertension, hypercholesterolemia, and metabolic disorders, remain among the leading contributors to the global burden of disease. These conditions generally develop gradually and are strongly influenced by dietary habits, physical activity, and lifestyle [1]. Consequently, health screening has become an essential preventive strategy for identifying health risk factors at an early stage, enabling individuals to receive appropriate health education and preventive interventions before their conditions worsen [2]. Recent advances in information technology have accelerated the digital transformation of healthcare services through web-based health information systems. These systems facilitate efficient collection, management, and retrieval of health examination records while improving data accessibility for both healthcare providers and the community [3], [4]. Furthermore, digital health information systems support the long-term storage of health records and improve the efficiency of health data management [5]. Nevertheless, most existing health screening systems remain limited to recording and presenting examination results without providing educational recommendations that help users interpret their health conditions and adopt healthier lifestyles [6].

*Corresponding author. *E-mail address: nasrul@nurulfikri.ac.id*
Received: 09 June 2026, Accepted: 30 July 2026 and available online 31 July 2026
DOI: https://doi.org/10.33751/komputasi.v23i2.108

Artificial Intelligence (AI), particularly Large Language Models (LLMs), has recently demonstrated considerable potential for supporting healthcare applications, including disease diagnosis, clinical decision support, medical text analysis, and health risk prediction [7], [8]. Despite these advances, most previous studies have focused primarily on clinical decision-making or diagnostic assistance. Only limited research has investigated the integration of LLM technology into health screening information systems for automatically generating educational fitness and nutrition recommendations based on categorized health screening results. Moreover, previous studies rarely discuss mechanisms to improve computational efficiency, recommendation consistency, and safe AI responses when integrating LLMs into web-based health information systems [9].

To address these limitations, this study proposes an AI-based health screening information system that automatically generates educational fitness and nutrition recommendations from categorized health screening results. The system processes several health indicators, including Body Mass Index (BMI), blood pressure, blood glucose, cholesterol, uric acid, and physical condition, to construct structured prompts that are submitted to an LLM through the OpenAI API. The generated recommendations are intended solely to support preventive healthcare and health education and are not designed to replace professional medical judgment or provide disease diagnosis.

The proposed system was developed using the XP software development methodology because XP supports iterative development, rapid adaptation to changing requirements, and continuous refinement throughout the implementation process [10]. Laravel was employed as the primary web application framework, while Filament was utilized as the administrative interface for efficient management of health screening data.

The primary novelty of this research lies in the proposed AI recommendation architecture rather than merely integrating an LLM into a web application. The proposed architecture combines structured prompt engineering, predefined AI guardrails, SHA-256-based prompt hashing, and recommendation caching within a web-based health screening information system. Structured prompt engineering ensures that health screening data are consistently transformed into AI-readable inputs, while guardrails constrain the model to generate educational recommendations without producing diagnoses, medication prescriptions, or unsupported assumptions. In addition, SHA-256 prompt hashing enables the system to identify identical requests and retrieve previously generated recommendations from the cache database instead of repeatedly invoking the AI service. By design, this mechanism is intended to reduce redundant API calls, lower computational overhead, shorten response time, and enhance recommendation consistency while maintaining the scalability of the proposed system; empirical benchmarking of these efficiency gains was beyond the scope of this study and is identified as a limitation.

The contributions of this study are threefold. First, it proposes an AI recommendation architecture that integrates structured prompt engineering, AI guardrails, prompt hashing, and recommendation caching into a health screening information system. Second, it demonstrates that the proposed architecture successfully satisfies all functional requirements through UAT. Third, it evaluates the stability of AI-generated recommendations through repeated-prompt consistency analysis, providing empirical evidence that the proposed recommendation mechanism produces reliable and reproducible outputs suitable for preventive health education.

2. Methods

This study adopted the Research and Development (R&D) methodology to design, implement, and evaluate an Artificial Intelligence-based health screening information system. R&D was selected because it supports the systematic development of software products while enabling iterative evaluation and refinement throughout the development process[11]. The proposed system was developed using the XP software development methodology due to its iterative nature, flexibility in accommodating changing requirements, and emphasis on continuous testing to improve software quality [12].

Figure 1 illustrates the overall research process adopted in this study.

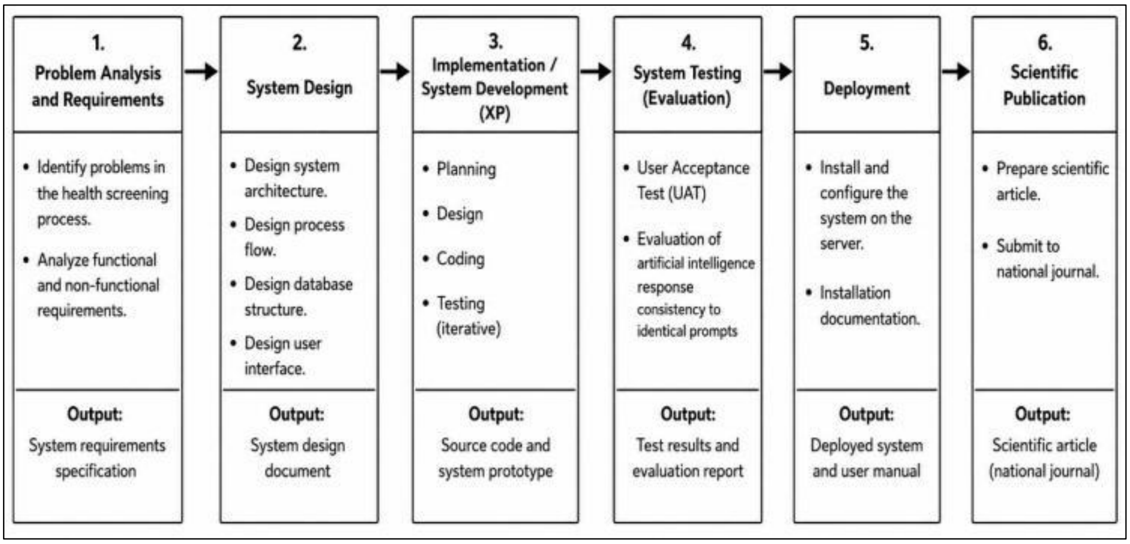

Figure 1. Research and system development stages.

The research consisted of four main phases: requirements analysis, system design, implementation, and evaluation.

During the requirements analysis phase, functional and non-functional requirements were identified through literature review and analysis of existing health screening workflows. The literature review focused on health screening information systems, Artificial Intelligence in healthcare, prompt engineering, Large language models (LLMs), and agile software development methodologies. Existing health screening procedures were also analyzed to determine the health parameters required for generating educational fitness and nutrition recommendations.

The system design phase focused on designing the software architecture, database schema, user interface, and Artificial Intelligence integration mechanism. The web application was developed using the Laravel framework, while Filament was employed as the administrative interface for managing users, health screening records, and recommendation data. Database entities were designed to store user profiles, health examination records, categorized screening results, generated recommendations, and prompt hashes used by the recommendation caching mechanism.

During the implementation phase, software development followed the XP lifecycle consisting of planning, design, coding, and testing [13]. The planning stage identified system requirements and AI recommendation features based on user needs. The design stage produced the application architecture, database model, user interface, and recommendation workflow. Coding was carried out using Laravel and Filament, while the Artificial Intelligence module was integrated through the OpenAI API. Throughout development, iterative refinement was applied to improve prompt engineering, guardrails, and recommendation generation based on evaluation results.

2.1. Proposed AI Recommendation Architecture

Figure 2 presents the architecture of the proposed AI-based health screening information system. The recommendation process begins when users submit health screening information through the web application. The collected parameters include Body Mass Index (BMI), blood pressure, blood glucose, cholesterol, uric acid, and physical condition. Rather than sending raw numerical measurements directly to the AI model, the system first converts each parameter into standardized health screening categories based on established clinical guidelines.

The categorized screening results are then combined with user profile information and transformed into a structured prompt using a predefined prompt engineering strategy. The prompt also includes guardrails that explicitly instruct the language model to generate educational recommendations only, prohibit medical diagnoses and medication prescriptions, and prevent unsupported assumptions beyond the available screening information.

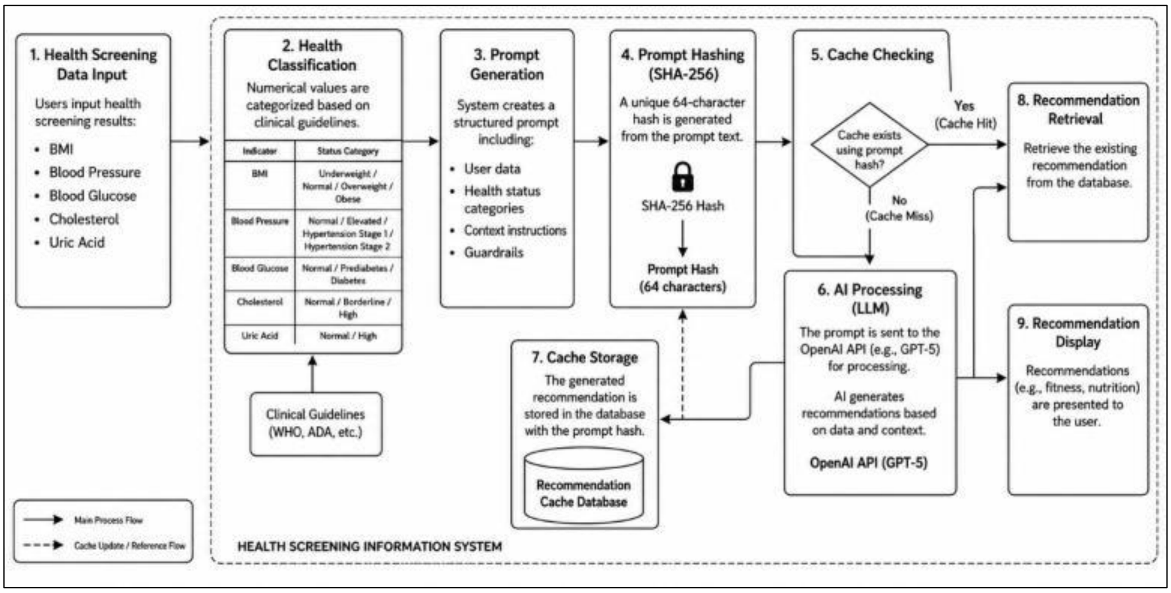

Figure 2. Architecture of the AI-Based Health Screening Information System with Prompt Hashing and Recommendation Caching

To improve computational efficiency, the generated prompt is converted into a SHA-256 hash value, which uniquely represents the complete prompt content. The system then searches the recommendation database using the generated hash.

When an identical prompt has previously been processed (cache hit), the corresponding recommendation is retrieved directly from the database without invoking the AI service. Otherwise (cache miss), the prompt is submitted to the GPT-5 Large Language Model through the OpenAI API. After the recommendation is generated, both the recommendation and its corresponding SHA-256 hash are stored in the database for future reuse.

Compared with conventional LLM integration that submits every request directly to the AI service, the proposed architecture is designed to reduce redundant API requests, shorten response time, lower operational costs, and improve recommendation consistency while maintaining scalability for web-based health screening applications.

2.2. Health Screening Classification

Before recommendations are generated, numerical health measurements are converted into standardized screening categories according to internationally recognized clinical guidelines. This classification process ensures that the AI model receives consistent health status descriptions instead of heterogeneous numerical values.

Table 1. Classification of Health Screening Parameters Based on Clinical Guidelines

Health Parameter	Clinical Guideline	Classification Used by AI
Body Mass Index (BMI)	WHO BMI Classification	Underweight, Normal, Overweight, Obese
Blood Pressure	ACC/AHA Hypertension Guideline	Normal, Elevated, Hypertension Stage 1, Stage 2
Blood Glucose	American Diabetes Association (ADA)	Normal, Prediabetes, Diabetes
Total Cholesterol	NCEP ATP III	Normal, Borderline High, High
Uric Acid	Indonesian Society of Internal Medicine / National Guideline	Low, Normal, High

As shown in Table 1, the classification stage standardizes heterogeneous clinical measurements before they are processed by the AI module. This approach simplifies prompt construction, improves recommendation consistency, and ensures that recommendations are generated according to established health screening criteria rather than arbitrary numerical values.

2.3. AI Recommendation Generation

After classification, the structured prompt is submitted to the GPT-5 Large Language Model through the OpenAI API. The model analyses only the categorized screening results provided by the system and generates educational recommendations without performing medical diagnosis or prescribing treatment.

The generated output consists of two principal recommendation categories:

1. Fitness Recommendations, containing physical activity guidance appropriate to the user's health screening category.
2. Nutrition Recommendations, containing dietary guidance and nutritional intake suggestions intended to support preventive healthcare.

The recommendations are presented through the web-based health screening information system in an understandable format, enabling users to interpret their health screening results and adopt healthier lifestyle practices while remaining aware that the recommendations serve only as educational support.

3. Result and Discussion

This study led to the development of a web-based health screening information system that leverages artificial intelligence to deliver automated fitness and nutrition recommendations. The application was built using the Laravel framework, with Filament serving as the administrative interface for data management, while recommendation generation was powered by a large language model (LLM). Through this integration, the system can evaluate users' health screening results and produce general fitness and nutrition recommendations based on standardized screening categories in an efficient and user-friendly manner.

3.1. Implementation of a Health Screening Information System

The health screening information system was developed as a web-based application using the Laravel and Filament frameworks. Filament serves as the admin panel for managing user data, health examination data, and AI-based recommendations. Filament simplifies interface development by providing modular features such as a form builder, table builder, relation manager, and statistics widget.

The proposed system integrates health data management with artificial intelligence to support recommendation services based on categorized health screening results. Its primary functions include managing user records, processing health examination results, generating AI-based recommendations, and presenting health insights. The evaluated health metrics comprise body mass index (BMI), blood pressure, blood glucose, cholesterol, and uric acid levels.

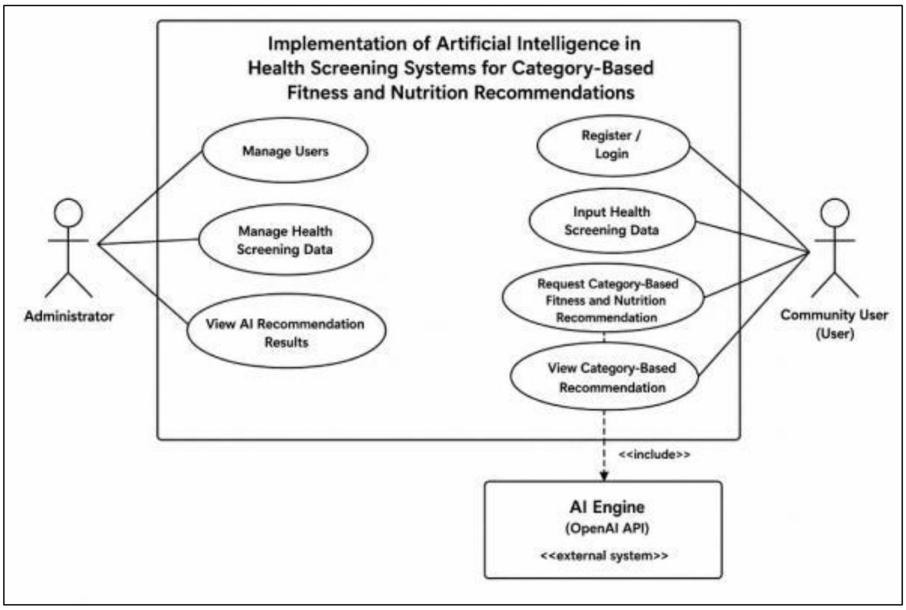

Figure 3. Use Case Diagram of the AI-Based Health Screening Information System

Figure 3 illustrates the interactions between administrators, community users, and the AI engine within the proposed health screening information system. Administrators manage user and health screening data, while community users can perform health examinations and obtain general fitness and nutrition recommendations based on their screening categories. The AI engine processes health screening data and generates recommendations based on users' health screening categories.

3.2. Implementation of Artificial Intelligence Integration

This study integrates AI into the health screening information system to automatically generate health screening-based fitness and nutrition recommendations. The AI module analyses categorized health screening data and produces educational recommendations intended to support preventive healthcare rather than replace professional medical judgment. The proposed system employs a GPT-5-based LLM accessed through the OpenAI API to analyze structured health screening information and generate contextual recommendations.

GPT-5 was selected because it achieved strong performance on the HealthBench Hard benchmark, which evaluates LLM capabilities in complex healthcare reasoning tasks.

The system uses prompt engineering to process users' health data. The parameters sent include BMI, physical condition, blood sugar, uric acid, cholesterol, and blood pressure. The system instructs the AI to generate four main components:

1. Health assessment.
2. Nutritional recommendations.
3. Physical fitness recommendations.
4. Healthy lifestyle advice.

The following prompts were used in this study. Algorithm 1. AI Prompt Construction:

```
Input:
    Kategori hasil skrining kesehatan pengguna (IMT, tekanan darah, gula darah,
    kolesterol, asam urat, dan kondisi fisik)
System Role:
    Anda adalah asisten edukasi kesehatan yang bertugas memberikan rekomendasi
    Kebugaran fisik dan gizi berdasarkan hasil skrining kesehatan pengguna.
Constraints:
    - Berikan rekomendasi hanya berdasarkan data skrining kesehatan.
    - Jangan memberikan diagnosis penyakit.
    - Jangan memberikan resep atau obat.
    - Jangan membuat asumsi di luar data yang tersedia.
    - Apabila terdapat parameter yang tidak normal, sarankan konsultasi dengan tenaga
      kesehatan.
    - Seluruh rekomendasi hanya bersifat edukatif dan sebagai pendukung pengambilan
      keputusan.
Output:
    1. Analisis Hasil Skrining
    2. Rekomendasi Kebugaran Fisik
    3. Rekomendasi Gizi
    4. Edukasi dan Tindak Lanjut
```

Algorithm 1 summarizes the prompt engineering strategy used in this study. Instead of directly sending raw health screening data to the LLM, the system constructs a structured prompt consisting of four components: the AI role, operational constraints (guardrails), categorized health screening data, and the expected output format. The prompt is intentionally written in Bahasa Indonesia because the target users are Indonesian citizens. This approach enables the AI to generate recommendations that are contextually appropriate, easy to understand, and consistent with the language used by the end users. The guardrails embedded in the prompt (Algorithm 1) function as the system's primary hallucination-mitigation mechanism: by restricting the model to reason only over the supplied categorized screening data, explicitly prohibiting disease diagnosis, medication prescriptions, and assumptions beyond the provided data, and requiring a referral to a healthcare professional whenever a parameter is abnormal, the prompt constrains the space of admissible outputs and reduces the likelihood of unsupported or fabricated clinical claims.

Algorithm 2 presents the implementation of the proposed SHA-256 prompt hashing mechanism.

```php
$promptHash = hash('sha256', $prompt);
$recommendation = RekomendasiAI::query()
    ->where('masyarakat_id', $masyarakat->id)
    ->where('prompt_hash', $promptHash)->latest()->first();
if ($recommendation) {
    return $recommendation->hasil_ai;
}
$response = Http::post($this->endpoint, $payload);
Recommendation::create(['prompt_hash'=>$promptHash,'hasil_ai'=>$response]);
return $response;
```

The generated recommendations are stored together with their SHA-256 prompt hashes, enabling the system to identify identical requests and reuse previously generated outputs. Compared with submitting every request directly to the AI service, this caching mechanism is expected to reduce response latency, minimize redundant API calls, and improve the scalability of the proposed health screening system, although this study did not perform quantitative server-load benchmarking to confirm these gains.

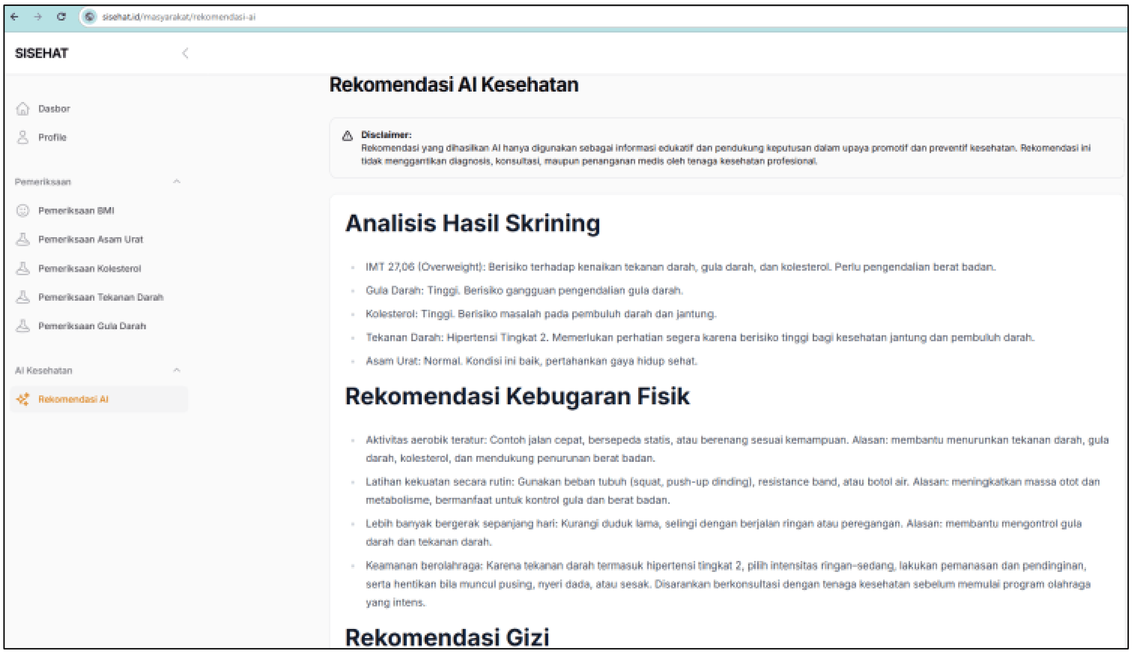

Figure 4. AI Recommendation Page

Figure 4 illustrates an example of AI-generated health recommendations based on categorized health screening results. The recommendations include health assessment, nutritional advice, physical fitness guidance, and preventive health education generated from the structured prompt.

3.3. Implementation of Extreme Programming

Figure 5 shows the implementation of the Extreme Programming (XP) method in the development of the proposed health screening information system. The XP methodology supports iterative software development through four main phases: planning, design, coding, and testing.

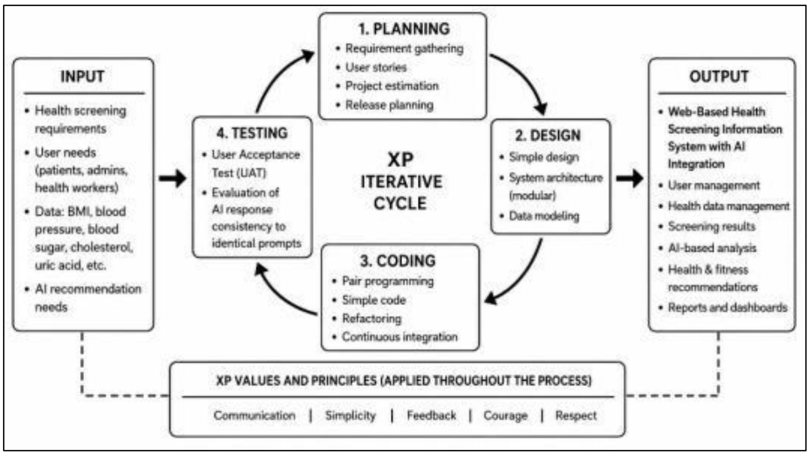

Figure 5. Implementation of Extreme Programming

The planning phase identifies the requirements for a health screening system, user needs, and AI-based health recommendation features [14]. The design phase includes the architecture, database, user interface, and mechanisms for integrating AI services [15]. System development was conducted using Laravel with Filament serving as the administrative interface. In addition, the OpenAI API was integrated to automatically produce fitness and dietary recommendations based on categorized health screening data [16].

3.4. Evalution

Testing was conducted iteratively throughout the software development process to ensure that the proposed system satisfied both functional requirements and artificial intelligence performance requirements. The evaluation framework consisted of three complementary components:

1. User Acceptance Testing (UAT), which verified whether the implemented system fulfilled all functional requirements.
2. Evaluation of artificial intelligence response consistency to identical prompts.

User Acceptance Testing (UAT)

User Acceptance Testing (UAT) was conducted to verify that all functional requirements identified during the analysis phase were successfully implemented. The evaluation covered authentication, user management, health screening data management, recommendation generation, recommendation storage, and AI service integration based on predefined testing scenarios [17].

The UAT results are presented in Table 2. Testing was conducted on all major system features, including user authentication, health examination data management, artificial intelligence integration, recommendation generation, recommendation storage, and recommendation presentation.

Table 2. User Acceptance Test Results

No	Feature Tested	Test Scenario	Expected Result	Test Result	Status
1	System Login	The user enters a valid username and password	The system successfully logs in to the dashboard	Correct	Valid
2	User Data Management	Add, edit, and delete user data	Data is stored and updated correctly	Correct	Valid
3	Health Examination Data	Add BMI, blood pressure, blood sugar, cholesterol, and uric acid data	Data is stored in the database	Correct	Valid
4	Examination History	Displays the user's examination history	The history is displayed based on the stored data	Correct	Valid
5	AI Integration	Sending health data to the AI service	Data successfully sent and processed	Correct	Valid
6	Recommendation Generation	Generating recommendations based on health data	The system generates automatic recommendations	Correct	Valid
7	Recommendation Storage	Save AI recommendation results to the database	Recommendation data is stored correctly	Correct	Valid
8	Recommendations View	Displays recommendation results to users	Recommendations are displayed properly	Correct	Valid
9	Data Search	Performing user data searches and checks	The data being searched for can be found	Correct	Valid
10	System Logout	The user logs out of the application	The system returns to the login page	Correct	Valid

As presented in Table 2, acceptance testing demonstrated that all predefined functional scenarios were successfully executed. All 10 test scenarios passed without failure, resulting in a functional success rate of 100%. These results indicate that the proposed system satisfies the specified functional requirements.

Evaluation of Artificial Intelligence Response Consistency to Identical Prompts

To evaluate the reliability of the proposed recommendation mechanism, an AI Recommendation Consistency Evaluation was performed using identical structured prompts generated from the same health screening records stored in the recommendation database. For each selected user, the same structured prompt was submitted independently to the AI model five consecutive times under identical experimental conditions.

The generated recommendations were analyzed using a mixed-methods approach combining qualitative content analysis with quantitative lexical similarity measurements. Qualitative analysis evaluated seven predefined dimensions, including recommendation structure, interpretation of health screening results, nutrition recommendations, aerobic exercise recommendations, strength training recommendations, health education and follow-up, and response completeness. Quantitative evaluation employed TF–IDF cosine similarity and Jaccard similarity to measure lexical consistency among generated recommendations. In addition, descriptive statistics, including the mean response length, standard deviation, and coefficient of variation (CV), were calculated to assess output variability [18]. Previous studies have demonstrated that lexical similarity metrics are effective for evaluating the reproducibility and stability of Large Language Model outputs across repeated generations [19].

To test the consistency (reliability) of a generative AI model's responses, one identical prompt was tested five times independently, producing five responses (R1–R5). The evaluation employed a mixed-methods approach combining:

1. qualitative deductive content analysis using seven a priori categories-structure/format, interpretation of screening results, aerobic exercise recommendations, strength training recommendations, nutrition recommendations, education/follow-up, and response length-scored for similarity (0–100%) manually.
2. TF-IDF cosine similarity, a measure of lexical similarity between documents based on Term Frequency-Inverse Document Frequency weighting, a standard technique in natural language processing (values 0–1)
3. Jaccard similarity to measure the overlap of unique vocabulary between responses
4. Descriptive statistics (mean, standard deviation, coefficient of variation/CV) on word count as an indicator of consistency in output depth. The triangulation of these four methods was intended so that the weakness of one method (e.g., the subjectivity of qualitative scoring) could be offset by other, more measurable methods.

Qualitative Evaluation Results

Table 3. Inter-Response Consistency Scores Based on Qualitative Content Analysis

Evaluation Dimension	Consistency Score	Remarks
Response structure and format	100%	Identical structure: Screening Analysis, Fitness Recommendations, Nutrition Recommendations, Education/Follow-up; presented in bullet-point format.
Interpretation of screening results	98%	Identical risk labels for each parameter (overweight BMI, high blood glucose, high cholesterol, stage 2 hypertension, normal uric acid).
Nutrition recommendations	92%	All responses use the "plate method" (1/2 vegetables, 1/4 protein, 1/4 carbohydrates) and restriction of sugar/salt/saturated fat; variation occurs in the depth of explanation.
Aerobic exercise recommendations	90%	All responses recommend 150–300 minutes/week, light–moderate intensity, with a gradual increase from 10–15 minutes/session.
Strength training recommendations	90%	All responses recommend 1–2 sets × 10–15 repetitions with light loads, 2–3 times/week, avoiding breath-holding.
Education and follow-up	88%	All responses recommend consultation with a healthcare provider and blood pressure monitoring; completeness of emergency warning signs varies slightly.
Length and depth of response	75%	Word count 553–727 (mean 617.6; SD 62.0; CV ≈ 10%); variation in elaboration despite consistent substance.

The average qualitative consistency score across the seven dimensions was 90.4%, classified as high consistency.

Quantitative Evaluation Results

Table 4. TF-IDF Cosine Similarity Matrix Between Responses

	R1	**R2**	**R3**	**R4**	**R5**
R1	1.000	0.823	0.777	0.845	0.786
R2	0.823	1.000	0.790	0.764	0.799
R3	0.777	0.790	1.000	0.747	0.772
R4	0.845	0.764	0.747	1.000	0.735
R5	0.786	0.799	0.772	0.735	1.000

The average pairwise cosine similarity between responses was 0.784 (SD = 0.031; range 0.735 - 0.845), indicating high and stable lexical similarity. As a comparison, the average Jaccard similarity (ratio of unique vocabulary intersection) was 0.478 - lower because this method is highly sensitive to diction variation resulting from paraphrasing, even when the meaning remains the same.

Table 5. Descriptive Statistics of Word Count per Response

Response	Word Count	Deviation from Mean
Response 1	553	-64.6
Response 2	642	+24.4
Response 3	727	+109.4
Response 4	590	-27.6
Response 5	576	-41.6

The average word count across the five responses was 617.6 (SD = 62.0; CV ≈ 10%). A CV below 15% indicates low-to-moderate variability in output length.

The integration of qualitative and quantitative results reveals a convergent pattern. At the level of substance-interpretation of screening results, physical activity recommendations, and nutrition recommendations-the five responses demonstrated very high consistency (88–98%), supported by a high and stable cosine similarity value (mean 0.784; SD 0.031). This indicates the model's reasoning stability with respect to the same prompt, without any contradictory or potentially medically harmful recommendations.

In contrast, the more moderate Jaccard similarity (0.478) and the variation in response length (CV ≈ 10%) indicate that inter-response variation occurred primarily at the surface level (diction, order of presentation, degree of elaboration) rather than in the substance of the recommendations—consistent with the probabilistic (non-deterministic) nature of generative language models. These findings indicate that the model is sufficiently reliable for similar use cases; however, the completeness of safety details (e.g., the list of emergency warning signs), which still varies, warrants attention, since in a health context this should ideally remain consistent across every output.

Scoring in the qualitative content analysis (Table 3) was conducted by a single rater, which carries a potential for subjective bias; further research is recommended to involve independent raters and to calculate inter-rater reliability (e.g., Cohen's Kappa/Krippendorff's Alpha). TF-IDF and Jaccard similarity also measure only lexical similarity, not semantic similarity; sentence-embedding-based measurement is recommended for further research to strengthen the validity of consistency claims at the level of meaning.

3.5. Discussion

The findings of this study demonstrate that integrating Artificial Intelligence into a web-based health screening information system extends its functionality beyond health data management to the automatic generation of educational fitness and nutrition recommendations. The proposed system combines categorized health screening data with prompt engineering, predefined guardrails, and a Large Language Model (LLM) to generate recommendations that are understandable, contextually appropriate, and intended to support preventive healthcare rather than replace professional medical judgment.

Unlike previous studies that primarily focused on disease diagnosis, medical data classification, or health risk prediction [7], [8]. This research integrates AI directly into a health screening information system through a structured recommendation generation mechanism. The proposed architecture incorporates health data categorization, prompt engineering, SHA-256 prompt hashing, recommendation caching, and OpenAI API integration into a unified framework. This architecture is designed to reduce redundant API requests, improve computational efficiency, and maintain recommendation consistency for identical health screening inputs while supporting scalable deployment in web-based health screening services.

The functional evaluation demonstrated that all major system features successfully passed the UAT, indicating that the proposed system satisfied the specified functional requirements and operated reliably across all tested scenarios. Furthermore, the qualitative content analysis conducted as part of the consistency evaluation confirmed that the generated recommendations were consistent with the categorized health screening results and fully complied with the predefined guardrails. The recommendations did not contain disease diagnoses, medication prescriptions, or unsupported assumptions, indicating that the prompt engineering strategy effectively constrained the Large Language Model to generate educational and preventive health recommendations appropriate for community health screening.

The AI Recommendation Consistency Evaluation further demonstrated that the proposed recommendation mechanism produced stable and reproducible outputs when identical health screening data were repeatedly processed. The qualitative analysis yielded an average consistency score of 90.4%, while the quantitative evaluation produced an average TF–IDF cosine similarity of 0.784, indicating a high degree of lexical consistency among repeated recommendations. Although variations in vocabulary and response length were observed, these differences did not affect the interpretation of health screening results or the substance of the generated recommendations. These findings indicate that the proposed prompt engineering strategy, together with the SHA-256 prompt hashing and recommendation caching mechanism, contributes to producing reliable and consistent AI-generated recommendations.

The primary contribution of this research lies in the development of an AI-based health screening information system that integrates structured prompt engineering, guardrails, recommendation caching, and Large Language Model technology to generate educational fitness and nutrition recommendations based on categorized health screening results. The proposed architecture not only supports the digitalization of health screening services but also improves computational efficiency and recommendation reliability through prompt hashing and recommendation reuse mechanisms.

Despite these contributions, this study has several limitations. The quality of the generated recommendations depends on the completeness and accuracy of the health screening data provided by users, while the evaluation focused on recommendation appropriateness and consistency rather than clinical effectiveness. In addition, the generated recommendations have not yet undergone comprehensive clinical validation by healthcare professionals. The system currently generates general recommendations based on categorized screening results (BMI, blood pressure, blood glucose, cholesterol, uric acid, and physical condition) and does not yet incorporate other clinically relevant variables such as age, sex, lifestyle factors, symptoms, medication use, existing diagnoses, or allergies; this study therefore reports general, category-based recommendations rather than fully individualized ones. Furthermore, the efficiency benefits attributed to the SHA-256 prompt hashing and recommendation caching mechanism (e.g., reduced server load, lower latency, and lower operational cost relative to a non-cached implementation) were not empirically benchmarked in this study. Future research may incorporate additional health variables, expert-based clinical validation, semantic similarity evaluation techniques based on sentence embeddings, domain-specific healthcare language models, and quantitative benchmarking of caching efficiency (e.g., server load and response latency with versus without caching) to further improve recommendation quality and reliability.

4. Conclusion

This study proposed and implemented an Artificial Intelligence-based health screening information system capable of automatically generating educational fitness and nutrition recommendations from categorized health screening results. The system integrates a Large Language Model (LLM) via the OpenAI API with structured prompt engineering, predefined guardrails, SHA-256 prompt hashing, and a recommendation caching mechanism to produce consistent, contextually appropriate recommendations while improving computational efficiency.

The evaluation results demonstrate that the proposed system successfully satisfied all functional requirements through User Acceptance Testing (UAT), indicating that all major system features operated correctly according to the specified requirements. In addition, the AI Recommendation Consistency Evaluation showed that repeated submissions of identical health screening data produced stable and reproducible recommendations. The qualitative analysis achieved an average consistency score of 90.4%, while the quantitative evaluation yielded an average TF–IDF cosine similarity of 0.784, indicating a high degree of consistency across repeated AI-generated recommendations. These findings demonstrate that the proposed prompt engineering strategy and recommendation generation mechanism can produce reliable educational recommendations while maintaining consistency for identical inputs.

The primary contribution of this research is the proposed architecture, which combines categorized health screening data, prompt engineering, AI guardrails, SHA-256 prompt hashing, recommendation caching, and LLM-based recommendation generation within a single web-based health screening information system. Compared with conventional AI integration approaches that directly invoke the language model for every request, the proposed mechanism improves scalability by reducing redundant API calls while maintaining recommendation consistency; empirical validation of these efficiency gains remains a direction for future work. Beyond the specific system implemented in this study, these findings carry a broader implication for software engineering practice with LLM-integrated systems: request-level hashing and response caching, combined with prompt-level guardrails, offer a general architectural pattern for constraining and stabilizing non-deterministic model outputs in production-grade information systems, independent of the particular domain of application.

Although the proposed system demonstrates promising results, several limitations remain. The generated recommendations depend on the completeness and accuracy of the health screening data provided by users, are based only on categorized screening status rather than a broader set of clinical variables (e.g., age, sex, lifestyle, symptoms, medication use, existing diagnoses, and allergies), and are intended solely as general educational support rather than personalized medical diagnosis or treatment. In addition, the anticipated efficiency gains of the SHA-256 prompt hashing and recommendation caching mechanism were not quantitatively benchmarked against a non-cached implementation. Future research should incorporate additional health variables, involve healthcare professionals in clinical validation, evaluate semantic consistency using embedding-based similarity methods, investigate domain-specific healthcare language models, and empirically benchmark server load and response latency with and without caching to further improve recommendation quality, personalization, and reliability.

5. Acknowledgement

The author gratefully acknowledges the financial support provided by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia through the Beginner Lecturer Research (PDP) grant program, which made this research possible. Appreciation is also extended to the Institute for Research and Community Service (LPPM) of STT Terpadu Nurul Fikri for its continuous assistance, valuable guidance, and institutional support during the completion of this study.

Special thanks are due to the research team members for their collaboration, dedication, and contributions to the development of the system and the completion of this study. The author also acknowledges all parties involved, including academic colleagues and participants, who provided valuable input and assistance during the system's implementation and testing.

It is hoped that this study will contribute to the ongoing development of AI-driven health information systems and encourage the delivery of more effective and accessible digital healthcare services for the community.

References

[1] M. Maesaroh, Y. Febsi, E. Z. V. Rinjani, J. Alfira, N. Noorlinda, and A. D. Rangga, "Penguatan Kesadaran Kesehatan Masyarakat Melalui Skrining Gula Darah Dan Tekanan Darah," J. Pengabdi. Masy. Bunda Delima, vol. 4, no. 2, pp. 138–149, Aug. 2025, doi: 10.59030/jpmbd.v4i2.104.

[2] A. Apriyana and N. Nasrul, "Pengembangan Sistem Skrining Kesehatan Berbasis Web dengan Metode Extreme Programming di STT-NF," DBESTI J. Digit. Bus. Technol. Innov., vol. 2, no. 2, pp. 215–223, Dec. 2025, doi: 10.54914/dbesti.v2i2.2022.

[3] F. Septiani Santoso, P. A. Ramadhani, D. Amnamuchlisah, and S. H. Purba, "Transformasi Digital Dalam Sektor Kesehatan Kajian Literatur Untuk Mendukung Inovasi dan Efisiensi Layanan Kesehatan," Cindoku J. Keperawatan dan Ilmu Kesehat., vol. 2, no. 1, pp. 1–12, Jan. 2025, doi: 10.61492/cindoku.v2i1.240.

[4] S. Cahyanti, T. P. Zahrani, B. R. Amalya, A. S. Cahyono, and S. Hariyanto, "Penerapan Sistem Informasi Manajemen Puskesmas (Simpus) dalam Meningkatkan Pelayanan Kesehatan di UPT Puskesmas Beji Tulungagung," J. Gov. Public Adm., vol. 2, no. 2, pp. 323–330, Mar. 2025, doi: 10.70248/jogapa.v2i2.1668.

[5] A. Tuzzahra, D. Aqila, R. S. Arenja, S. Naifah, N. A. Rinovi, and M. Sukma, "Transformasi Digital Kesehatan di Indonesia: Analisis Komprehensif Peran dan Tantangan Rekam Medis Elektronik," J. Teknol. Sist. Inf., vol. 7, no. 1, pp. 59–68, May 2026, doi: 10.35957/jtsi.v7i1.14333.

[6] S. Kusumadewi, R. Kurniawan, E. Marfianti, and A. Khodzim, "Edukasi Masyarakat Tentang Pemanfaatan Aplikasi Kesehatan Online Melalui Partisipasi Pengembangan Sistem Informasi," JPPM (Jurnal Pengabdi. dan Pemberdaya. Masyarakat), vol. 4, no. 2, p. 337, Apr. 2021, doi: 10.30595/jppm.v4i2.8597.

[7] D. L. Dede, Subhiyanto, Esthi Adityarini, and Mochamad Arief Madiansah, "Analisis Implementasi Kecerdasan Buatan (Artificial Intelligence) Dalam Optimalisasi Proses Bisnis," J. Sist. Inf. dan Teknol., vol. 5, no. 1, pp. 90–99, Jan. 2025, doi: 10.56995/sintek.v5i1.135.

[8] R. Ardhi Harlanto, D. Iskandar, and C. Ajika Pamungkas, "Pemantauan Kesehatan Berbasis Kecerdasan Buatan untuk Perekaman dan Analisis Histori Kesehatan," JATI (Jurnal Mhs. Tek. Inform., vol. 9, no. 4, pp. 7252–7257, May 2025, doi: 10.36040/jati.v9i4.14414.

[9] A. Wisni et al., "Skrining kesehatan dasar dan edukasi berbasis komunitas," BEMAS J. Bermasyarakat, vol. 6, no. 2, pp. 313–324, Oct. 2025, doi: 10.37373/bemas.v6i2.1826.

[10] Nasrul, Edi Wibowo, and David Wahyu Wismanindra, "Design and Implementation of a Teaching Assistant Information System Using Laravel Filament and Extreme Programming," Komputasi J. Ilm. Ilmu Komput. dan Mat., vol. 23, no. 1, pp. 89–99, Jan. 2026, doi: 10.33751/komputasi.v23i1.83.

[11] Zulkarnaini, M. Noval, Andre Mariza Putra, and A. Octarina, "Pengembangan Sistem Informasi P3m Terintegrasi Melalui Refactoring dan Penambahan Fitur dengan Metode R&D," Rabit J. Teknol. dan Sist. Inf. Univrab, vol. 11, no. 1, pp. 1546–1558, Jan. 2026, doi: 10.36341/rabit.v11i1.7268.

[12] N. Nasrul, H. Saptono, E. Wibowo, and A. Amalia, "Rancang Bangun Sistem Informasi Manajemen Aset Berbasis Web untuk Menghitung Penyusutan Fiskal," J. Inform. Terpadu, vol. 10, no. 1, pp. 66–72, Mar. 2024, doi: 10.54914/jit.v10i1.1086.

[13] J. Junaida, L. Zehro, S. H. U. Hasanah, and A. Hamdani, "Desain Sistem Informasi Deteksi Dini Penyakit Tidak Menular (Obesitas) Berbasis Web Di Departemen Kesehatan Banyuwangi," J. Ilm. Sist. Inf., vol. 5, no. 2, pp. 130–140, Jan. 2026, doi: 10.51903/7zcp9458.

[14] H. J. Reppi, H. K. Manggopa, and D. R. Kaparang, "Pengembangan Sistem Informasi Tata Usaha Menggunakan Metode Extreme Programming (Studi Kasus: Jurusan Pendidikan Teknologi Informasi dan Komunikasi)," Edutik J. Pendidik. Teknol. Inf. dan Komun., vol. 5, no. 4, pp. 945–959, Jul. 2025, doi: 10.53682/d8v8th72.

[15] D. V. Exanova, N. S. -, and B. I. H. -, “Rancang Bangun Portal Transparansi Anggaran Berbasis Website pada Tim Pembangunan Telaga Harapan,” J. Inform. dan Tek. Elektro Terap., vol. 14, no. 2, Apr. 2026, doi: 10.23960/jitet.v14i2.9321.

[16] F. M. J. Kanoko, M. Adi Paramartha Putra, and A. A. Gede Adi Mega Putra, “Rancang Bangun Website Untuk Pengelolaan Informasi Dan Promosi Desa Wisata Jehem,” JATI (Jurnal Mhs. Tek. Inform., vol. 10, no. 2, pp. 3575–3579, Mar. 2026, doi: 10.36040/jati.v10i2.17905.

[17] A. Nur Sahid, Y. Agustiansyah, and Apriadi, “Penerapan Metode Extreme Programming dalam Pengembangan Sistem Manajemen Proyek untuk Meningkatkan Efisiensi Kerja,” J. Technol. Appl. Sci., vol. 1, no. 1, pp. 21–32, Jan. 2026, doi: 10.64878/jtas.v1i1.87.

[18] K. Sato, “Exploring the Educational Landscape of AI: Large Language Models’ Approaches to Explaining Conservation of Momentum in Physics,” 2024, [Online]. Available: http://arxiv.org/abs/2407.05308

[19] K. Lee, “Consistency of AI-Generated Exercise Prescriptions : A Repeated Generation Study Using a Large Language Model”.