Study on an Intelligent Review System for Online Health Information Queries by the Elderly

Ying Wang, Yutong Li, Shaohan Chen, Mei Wang*
College of Medical Information and Artificial Intelligence, Shandong First Medical University and Shandong Academy of Medical Sciences, Tai'an, Shandong, China
**Corresponding Author.*

Abstract: The Internet has become the main channel for Chinese elderly netizens to obtain health information. But these elderly people have low digital literacy. In addition, the health information on the network is both correct and wrong. And many websites are not designed for the elderly. So these problems may affect their health and use up unnecessary medical resources. To solve these problems, we have developed and put into practice the 'Silver Smart Care' Intelligent Review System. This system has four parts that work together in a loop. These components include Azure Artificial Intelligence (AI) for problem discovery, Transformer Bidirectional Encoder Representation (BERT) for semantic analysis, Long Short-Term Memory (LSTM) for trend prediction, and an interface designed for older users. Therefore, the system can filter information, identify abnormal queries, provide correct medical advice, and share personalized health education content. Experimental results show that the system can find 91.3% of abnormal queries and 89.7% of the health information. For the elderly users, the efficiency in completing tasks improves to 80.5%, the time spent searching for information is reduced and the interface satisfaction reaches 9 out of 10.

Keywords: Health Information Review; Anomaly Detection; Age-friendly Design; BERT; LSTM

1. Introduction

By 2024, China's population aged 60 and over will reach 282 million, accounting for 19.9 % of the total population; the number of Internet users aged 60 and above reached 158 million, accounting for 14.4 % of the total number of Internet users. The Internet has become the primary channel for the elderly to obtain health information and consult health issues. However, relevant data show that only 33.6 % of elderly netizens can effectively evaluate the credibility of online health information, while 41.5 % of elderly netizens are at risk of misinterpreting such information. Misleading false health information and pseudo-scientific content can easily lead to the misuse of drugs or medical treatment in the absence of professional guidance for the elderly, which not only threatens their health, but also causes waste of medical resources.

At present, there are three main problems in the online health information service for the elderly: first, the information source is messy, and the lack of effective audit mechanism leads to the proliferation of false advertising and pseudo-scientific content; second, the platform has insufficient adaptability to elderly users, and the font is too small and the operation is complex, which does not conform to the physiological characteristics of the elderly. Third, there is a lack of targeted behavior monitoring and risk early warning mechanisms, and abnormal search behaviors cannot be intervened in time. The State Council's "Implementation Plan for Effectively Solving the Difficult Problem of Using Intelligent Technology for the Elderly" clearly requires promoting the aging transformation of Internet applications, strengthening the protection of health information services, and bridging the digital divide of the elderly. Although there is clear policy guidance, most of the existing research focuses on optimizing a single module, or fails to fully integrate the specific needs of the elderly, resulting in a lack of comprehensive solutions that combine review, monitoring, services and public education. Based on this, this study combines medical information engineering with artificial intelligence technology to develop an intelligent review system to solve the problems faced by the elderly in obtaining online health

information, and to support the implementation of healthy China and smart aging strategies.

2. Current State of Research at Home and Abroad

2.1 Current State of Research Abroad

2.1.1Recent research developments

Developed countries such as Europe, the United States and Japan have a high degree of population aging, the application of digital health technology is relatively mature, and a relatively complete health information service system for the elderly has been established. Foreign research mainly focuses on the health information retrieval behavior of elderly users, information credibility evaluation and intelligent technology auxiliary services. Some studies have pointed out that the elderly group faces problems such as weak discrimination ability and susceptibility to false information when searching for health information on the Internet. Therefore, it is necessary to improve the efficiency of information verification and service security through technical means.[1] Some developed countries have cooperated with authoritative medical institutions to establish online health service platforms, use artificial intelligence algorithms to verify the authenticity of health information, and adopt aging-friendly interaction design to reduce the use threshold of elderly users. [2] In terms of data governance, the European Union 's General Data Protection Regulation (GDPR) and the United States ' Health Insurance Circulation and Liability Act (HIPAA) and other regulations strictly regulate the protection of health data privacy from a legal perspective, and establish a mature ethical and security framework.[3]

2.1.2 Limitations of the research

Although significant progress has been made in international research and practice, these systems are designed based on their respective national healthcare systems and user habits, resulting in significant differences from China in terms of business processes, interaction logic, and linguistic and cultural contexts. For example, their interaction designs often focus on the family doctor referral model, which differs from the behavioral habits of elderly users in China who tend to seek medical care directly. Consequently, these systems are difficult to directly adapt to the domestic healthcare context and the characteristics of the elderly population [4].

2.2 Current State of Domestic Research

2.2.1 Progress in existing research

The 54th statistical report of China 's Internet development points out that the number of elderly Internet users in China continues to grow, but the overall ability to identify health information is still weak, and false health information has become a hidden danger threatening the health of the elderly [5]. In terms of technical research, The deep learning method is used to realize the automatic identification of false health information on the Internet, which provides a technical basis for health information verification. Hou Guanhua et al.carried out research on the interactive design of health information system for the elderly according to the characteristics of elderly users, so as to improve the usability of the system[6]. Relevant scholars have also realized the health risk prediction of the elderly based on LSTM, which provides support for the analysis of time series health data.

2.2.2 Limitations of the research

Although the existing research covers the fields of health information review, age-appropriate design and health prediction, there are still significant deficiencies. Health information review often relies on keyword matching, resulting in insufficient recognition accuracy at the semantic level. Most of the elderly-oriented transformation is limited to interface adjustment, failing to adapt to the physiological and cognitive characteristics of the elderly in terms of process and interaction logic. There is a lack of monitoring mechanism for user search behavior and early warning mechanism for abnormal situations. Most systems only optimize a single function, and have not yet established a service system covering behavior monitoring, intelligent review, precision medical guidance and health education push. It is difficult to systematically solve the challenges faced by the elderly in accessing health information [7].

2.3 Approach to the Research

In summary, there are still some shortcomings in the current research, such as scattered health information review, lack of early warning system for abnormal search behavior, superficial adaptation for the elderly, disjointed service chain and insufficient localization. At the same time, the number of Internet users aged 60 and

above in China has reached 158 million. However, only 33.6 % of the elderly can effectively identify the authenticity of online health information, while 41.5 % of the elderly have the risk of misunderstanding information, and the situation is grim. This study addresses the full spectrum of pain points faced by older adults when accessing online health information—including a tendency to blindly trust folk remedies, frequent anxiety-driven searches, being misled by false health advertisements, and abandoning use due to complex operations. It systematically integrates multidisciplinary theories and methods from artificial intelligence, medical information engineering, and age-friendly interaction design, while deeply incorporating key technologies such as anomaly detection, BERT-based deep semantic understanding, LSTM temporal prediction, and collaborative filtering recommendations.

Most of the existing platforms only focus on content screening rather than user behavior, rely on keyword filtering that lacks medical semantic understanding, only expand the text without reconstructing the workflow, and push common content instead of customized according to personal health needs. In order to solve this defect, this study makes improvements in the following three aspects: First, for the common abnormal behaviors of the elderly (such as repeatedly searching for high-risk content or frequently querying critical symptoms), the anomaly detection model is used to achieve real-time identification and active intervention. Secondly, in view of the auditing problems such as pseudoscience and false medical advertisements, BERT deep semantic understanding algorithm is used to accurately determine the source credibility, scientific validity and evidence chain, timeliness and other dimensions, which effectively solves the problem of low detection rate of false information at the semantic level inherent in traditional keyword filtering [8]. Thirdly, based on the application of LSTM in health risk prediction and time series analysis of the elderly [9], the prediction of health needs is realized, and personalized scientific communication content is accurately pushed. In addition, the whole user operation process including font, color matching scheme and operation process is optimized according to the interaction design standard for the elderly, which effectively reduces the operation threshold of the elderly users.

This study developed a closed-loop intelligent review system, which integrates real-time monitoring of search behavior, artificial intelligence-based health information review, accurate intelligent medical guidance, and personalized health education content push. Promote from passive information display to active risk intervention, from single function optimization to end-to-end service integration, and from general design to deep aging. Thus, a comprehensive solution covering behavior, information, service and health education is formed, which fills the research gap of the lack of integrated, closed-loop and implementable intelligent system in the field of elderly health information service. The system provides technical support and practical paradigm for bridging the digital divide of the elderly population, reducing the risk of information misreading, and promoting the Smart Health and Endowment (SHEC) plan and the Healthy China strategy[10].

3. System Overview

3.1 Design Objectives

Considering the challenges faced by the elderly when accessing health information online and their needs for age-appropriate design, four key design goals of the system were identified :

Integrate the resources of certified medical institutions and medical guidelines to build an authoritative medical knowledge base. Establish a multi-dimensional verification system to accurately identify false information and ensure the authenticity of information.

Real-time monitoring of search behavior to identify abnormal search patterns, risk warning and timely intervention to prevent the elderly from being misled by false information, so as to solve the problem of high risk of misjudgment of the elderly mentioned in the introduction.

Optimize the interface and interaction and design for the elderly, make up for the deficiency of the platform in the function of the elderly, and reduce the entry threshold of the elderly users.

Combine user portraits with artificial intelligence-driven medical guidance and personalized health education to provide targeted services, improve the efficiency of health information acquisition, and improve

health literacy.

3.2 System Architecture

As shown in Figure 1, the system adopts a three-tier architecture, and the functions of each layer are as follows:

As the cornerstone of the system, the data layer integrates four core data resources, including elderly user behavior data, medical knowledge base, popular science resource base (multi-modal content for the elderly) and user portrait database (including age, chronic disease status and search preferences), to provide data support for the realization of the upper functions.

As the core technology pillar, the algorithm layer integrates Azure AD, BERT semantic understanding, LSTM time series prediction and collaborative filtering recommendation algorithms to realize anomaly detection, information audit, demand prediction and content recommendation functions respectively.

The application layer is designed for elderly users, including four core functional modules: query behavior monitoring, AI-driven content review, precision medical guidance, and popular science content push. These modules correspond to the complete service workflow of monitoring, auditing, medical guidance and science popularization, and integrate the interactive components for the elderly to improve the user experience.

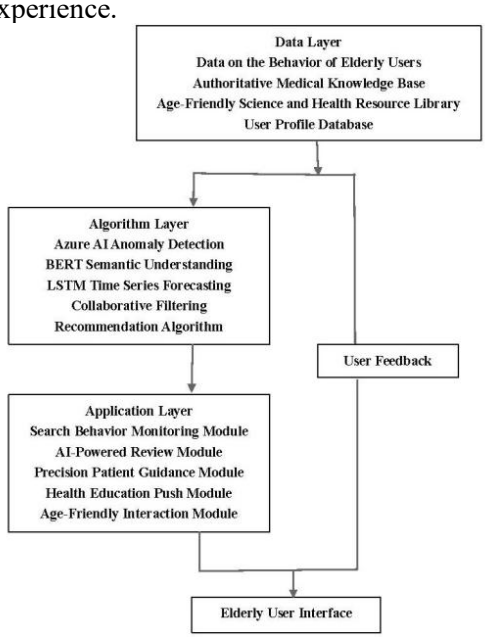

Figure 1. System Architecture Diagram

3.3 Core Module Design

3.3.1 Query behavior monitoring module

In order to cope with the abnormal query behavior patterns in elderly users, such as high-frequency repeated queries and search for folk bias. This module uses Azure AI anomaly detector to collect time series features to construct a baseline model of normal user behavior. Including query keywords, query frequency, browsing time and repetition rate. By comparing the degree of deviation between the current behavior and the benchmark model, abnormal patterns can be identified, including high-frequency queries for the same symptom in a short time, continuous search for folk bias, and repeated queries for high-risk symptoms. When abnormal behavior is detected, the system will trigger an early warning mechanism to actively intervene through voice prompts and pop-up alerts. The accuracy of anomaly detection is 91.3 %, which effectively reduces the risk of misleading information.

3.3.2 AI-Powered content moderation module

In order to solve the problem of uneven authenticity of network health information, this module integrates BERT semantic understanding algorithm to construct a medical terminology dictionary and a database of prohibited words. The system conducts a comprehensive and multi-dimensional assessment of online health information, including title, text, source reliability, evidence chain integrity and timeliness, and divides the information into three categories : real, suspicious and false. False information will be immediately blocked, suspicious information will be marked with warnings, and real information will be displayed normally. With 89.7 % audit accuracy, the system ensures the reliability of users ' access to information.

3.3.3 Precision medical guidance module

The module is designed to meet the specific medical needs of the elderly. It combines BERT semantic understanding and deep learning technology with medical knowledge base to realize the functions of symptom semantic analysis, health risk assessment, intelligent department recommendation, multi-round dialogue clarification and medical advice generation. It also supports voice and text input, which simplifies the user experience. With 87.5 % department recommendation accuracy, it ensures that the elderly can easily access medical guidance services.

3.3.4 Science popularization push module and age-friendly interaction module

The popular science push module uses user

portraits and LSTM time series prediction algorithms to predict the health needs of users, so as to push multimodal popular science content such as text, pictures, audio and video. The content adopts dialogue style and short text form to improve the understanding of elderly users, and the push relevance is ≥ 85 %.

The old-fashioned interactive module meets the WCAG AA barrier-free standard, adopts a large font design, the text text is not less than 16px, the title is not less than 24px, and adopts a high-contrast black and white color scheme. It also simplifies the operation process, supports voice input, text-to-speech, one-click query and historical record storage, and is compatible with mainstream screen readers, effectively reducing the threshold of use for the elderly.

4. Key Technical Implementation

4.1 Anomaly Detection Based on Azure AI

In this study, the time series data of search behavior of elderly users are systematically processed by feature engineering. Firstly, the core behavioral characteristics such as search frequency, browsing time, repetition times and search content categories are extracted, and the missing data are interpolated, and the boxplot is used to identify and process outliers. On this basis, by combining the isolated forest algorithm with the time series anomaly detection model, a user behavior anomaly detection model is constructed. The Azure AI anomaly detector monitors the real-time user behavior flow and calculates the anomaly score. When the score exceeds the preset threshold, the system will automatically trigger the alarm, and finally establish the processing mechanism of monitoring, marking, early warning and intervention. In the process of model training, grid search is used to optimize the key parameters of the isolated forest algorithm, such as the number of sub trees and the sampling rate, so as to balance the detection sensitivity and false alarm rate. The accuracy of anomaly recognition is 91.3 %.

4.2 Semantic Review of Health Information Based on BERT

In order to identify false health information through semantics, this study constructed a special dictionary containing medical terms, standardized expressions and common prohibited words, as well as a database of prohibited words, and integrated the authoritative medical knowledge map with the false information feature database. After preprocessing the online health information text, such as word segmentation, stop word removal and entity recognition, the data is input into the pre-trained BERT model for semantic coding. The model evaluates information from four dimensions: source credibility, scientific accuracy, compliance and timeliness, and implements hierarchical review based on preset thresholds. On the test set, the model achieves a high recall rate for false information, and the review accuracy rate is 89.7 %.

4.3 Health Needs Prediction and Personalized Recommendations Based on LSTM

This study collected the time series data of elderly users, including historical search records, browsing behavior and collection content, and combined these data with static features such as age, chronic disease history and health concerns to construct user behavior sequences. The LSTM time series prediction model is used to capture the time dependence in these behavior sequences and predict the user's potential health needs. On this basis, the collaborative filtering recommendation algorithm is used to match the predicted demand with the content in the popular science resource database to achieve personalized popular science recommendation. The recommended content is tailored according to the user 's portrait, and is consistent with the characteristics of the target user in terms of content theme, presentation form and language difficulty, with a matching rate of ≥ 85 %. This effectively improves the effectiveness of health education and helps elderly users improve their health literacy.

4.4 Age-Friendly Interface and Interaction Optimization Technologies

The system adopts a responsive front-end design, which can automatically adjust the interface layout and font size according to the screen size of the device, and is suitable for personal computers and Android mobile devices. It integrates speech recognition and text-to-speech technology, supports speech search and speech result broadcast, and solves operational problems such as typing difficulties and poor eyesight in the elderly. In terms of workflow, the system implements simplified navigation paths such as one-click search and one-click

consultation guidance. It retains historical search records and provides guidance prompts at key operational nodes to reduce cognitive load and minimize user error rates. The schematic diagram is shown in Figure 2.

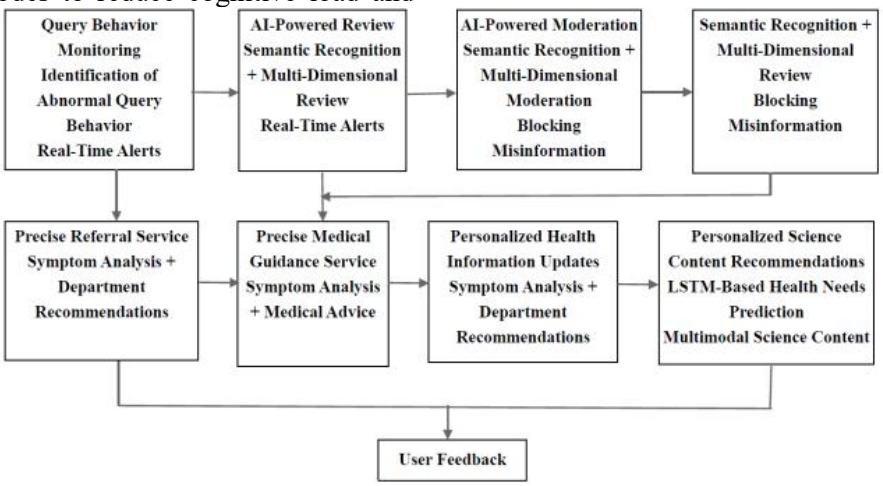

Figure 2. Schematic Diagram of a Closed-Loop Smart Health Service Process for Elderly Users

5. System Testing and Results Analysis

5.1 Test Environment

This study established the following test environment to ensure the authenticity and reliability of the test results.

Hardware environment: PC (Intel Core i5-12400F processor, 16GB memory, 512GB solid state drive), Android mobile devices (Huawei Mate 50,6.7-inch screen).

Software environment: Python 3.9 development language, MySQL 8.0 database management system, Azure AI cloud service platform, BERT pre-training model (bert-base-chinese) and Apache JMeter 5.5 performance test tool.

Test subjects: A total of 30 elderly users aged 60 and above were recruited, covering different age groups (15 60-69 years old, 10 70-79 years old, 5 80 years old and above). Subjects were grouped according to educational background (8 primary school and below, 12 junior high school, 10 high school and above), health status (20 with chronic diseases, 10 healthy) and digital literacy (8 low-level, 14 medium-level, 8 high-level) to ensure that the test samples were well representative.

5.2 Evaluation Metrics

According to the system design goals, three types of indicators are evaluated.

Functional performance indicators: abnormal query recognition accuracy, health information audit accuracy, AI recommended department matching rate and system response time.

User experience indicators: task completion rate, information retrieval time, interface satisfaction and error rate.

Health literacy indicators: health knowledge retention rate and ability to identify wrong information.

5.3 Testing Process and Results

The test is divided into three stages: function test, performance test and user experience test. Functional testing verifies whether the implementation of each core module meets the design expectations. Performance testing uses JMeter to simulate multi-user concurrent scenarios to evaluate system response speed and concurrent load capacity. The user experience test invited 30 test subjects to complete the specified task, recorded the task completion status, time consumption and error times during the whole process, and issued a satisfaction questionnaire after the task was completed.

The test results show that the accuracy of abnormal query recognition is 91.3 %, the accuracy of health information review is 89.7 %, and the AI-guided department recommendation matching rate is 87.5 %. The average response time of the PC is less than 0.5 seconds, and the mobile terminal is less than 0.8 seconds. The system runs stably under 50 concurrent users, and all functions operate normally.Compared with the user experience index before the system goes online, the user task completion rate increased from 28 % to 80.5 %, the information retrieval time shortened from 12 minutes to 6.4 minutes, the interface satisfaction score increased from 4.2 points to 9.0 points (full score of 10 points), and the error rate decreased

from 35 % to 7.8 %. In terms of health literacy indicators, the retention rate of health knowledge of elderly users increased by 38.2 %, and the ability to identify misinformation increased by 41.5 %. The three indicators have reached the design expectations.

6. Challenges and Prospects

At present, there are still some limitations in this study. The medical knowledge base has not fully covered rare diseases and some special diseases, resulting in a gap between the system function and the diversified needs of the elderly in chronic disease management and complex disease consultation. Although the current speech recognition function performs well in Mandarin, it lacks support for regional dialects, which may affect the experience of some elderly users. At present, this study has only been tested and verified in the laboratory environment, and has not yet carried out large-scale community applications, lacking long-term use of tracking data.

Future research will integrate the data resources of more authoritative medical institutions, continuously expand the medical knowledge base, and supplement the relevant content of rare diseases and specialist diseases to better meet the diverse health needs of the elderly population. Further optimize the speech recognition model, increase support for the main population dialects, and improve the convenience and universality of interaction. It is planned to cooperate with community health service centers and pension institutions to promote the system, collect long-term data in real use scenarios, and optimize system performance through continuous iteration. Explore the integration with the hospital system, build a service process covering consultation, triage and appointment, and improve service continuity and application value.

7. Conclusion

Aiming at the needs of the elderly to query health information online, an intelligent audit system is designed and implemented. By integrating core technologies such as Azure AD and BERT semantic understanding, a system framework integrating monitoring, auditing, referral and health education is constructed. The framework focuses on the key challenges faced by the elderly, and solves the problems of high risk of information misunderstanding, unfriendly platform for the elderly and the proliferation of false information. The test results of the system show that all the performance indicators have reached the design goals, which can effectively identify false health information and remind users to pay attention to abnormal search behavior, and significantly improve the experience and health literacy of elderly users. The research results have both theoretical significance and practical value. It not only enriches the academic research in the field of health information services for the elderly, but also provides a practical paradigm for the development of the smart health care (SHEC) industry. At the same time, it provides strong technical support for coping with the challenges of population aging and promoting the construction of' Healthy China.

Acknowledgments

This work was supported by Funding for the Undergraduate Innovation and Entrepreneurship Training Program (XJ2025104390447), Shandong First Medical University and Shandong Academy of Medical Sciences.

References

[1] China Internet Network Information Center. (2024). The 54th Statistical Report on the Development of the Internet in China. Beijing.

[2] Wang Sufang, Zhou Yuhan. (2022). A Study on the Ability of Older Adults to Evaluate Health Information on Social Media and Factors Influencing Its Improvement. Library and Information Service. 66(24):14-29.

[3] European Commission. (2022). Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space (EHDS). Brussels.

[4] Hu Chuntao, Qin Jinkang, Chen Jingmei, Zhang Liang. (2023). Deep Learning-Based Identification of Misinformation in Health Information. Computer Engineering and Applications. 59(15):234-240.

[5] Deng S L, Xia S D. (2022). Research on the Current Status and Influencing Factors of Older Adults' Ability to Discriminate Online Health Information. Library and Information Service. 66(10):89-98.

[6] Hou Guanhua, Liu Qiong. (2025). A Study on Age-Friendly Guided Design of Information Representation in Online

Health Information Systems. Journal of Packaging Science. 17(1):78-86.

[7] Li Jian, Zhang Yansong. (2024). Research on the Development of SHEC. Dalian: Northeast University of Finance and Economics Press.

[8] Yu Zhangxian, Mao Yuqing, Hu Kongfa. (2020). Deep Learning-Based Identification of False Health Information. Software Guide. 19(3):1-7.

[9] Wang Ling, Chen Shouqiang. (2023). LSTM-Based Prediction and Risk Early Warning of Elderly Health Time-Series Data. Chinese Journal of Digital Medicine. 18(5):45-49.

[10] Wang Weixuan, Shang Shaomei. (2024). Reflection on High-Quality Development of Smart Health Data. Strategic Study of CAE, 26(6): 32-42.

