Open Journal of Social Sciences, 2024, 12, 487-504
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

Study on the Analysis of Differentiated Demand for Citizen Healthcare Services Based on Data Mining

Hao Wan, Yihao Yin, Xin Li, Jing Yang

School of Management, Tianjin University of Technology, Tianjin, China
Email: wannahow0102@gmail.com

How to cite this paper: Wan, H., Yin, Y. H., Li, X., & Yang, J. (2024). Study on the Analysis of Differentiated Demand for Citizen Healthcare Services Based on Data Mining. *Open Journal of Social Sciences, 12,* 487-504.
https://doi.org/10.4236/jss.2024.127035

Received: June 12, 2024
Accepted: July 26, 2024
Published: July 29, 2024

Abstract

This study uses data mining techniques to deeply analyze the demand for healthcare services of citizens of different age groups in Tianjin to provide reliable data support for improving the quality and efficiency of healthcare services. By collecting data from a wide range of samples, including residents of different age groups and different social backgrounds, we applied advanced methods such as random forests to conduct a comprehensive analysis. It was found that there were significant differences in the demand for healthcare services among different age groups and social groups. Specifically, age, education level, and health status significantly impact the demand for healthcare services. Older age groups may be more inclined to seek long-term care and chronic disease management services, while younger people may be more concerned about emergency medical treatment and preventive care. In addition, our study found that more educated people may be more focused on the level of healthcare technology and service experience, while individuals with poorer health status may have a more urgent need for healthcare resources. These findings provide valuable references for healthcare organizations and policymakers, helping them to meet the healthcare service needs of different groups more accurately, improve the efficiency of healthcare resource utilization, and achieve the goal of health for all. Future research can further explore the mechanisms by which different factors influence the demand for healthcare services and make targeted policy recommendations to promote the continuous optimization and innovation of the healthcare service system.

Keywords

Data Mining, Healthcare Services, Differentiated Demand, Random Forests

1. Introduction

With the rapid development of medical technology and the public's growing concern for health, medical services have become an indispensable part of social life. However, demographic changes, lifestyle diversification, and an increase in chronic diseases have made citizens' needs for medical services increasingly diverse. For instance, in China, the demand for geriatric medical services is gradually increasing with the aging population, and the proportion of elderly people suffering from chronic diseases is on the rise. This requires medical institutions to better understand and meet the needs of patients of different ages and health conditions by providing personalized and differentiated medical services. In such a context, the application of data mining techniques becomes particularly important.

In recent years, big data technology has significantly impacted various industries, including healthcare (National Health and Family Planning Commission, 2017). The U.S. government's investment in the "Big Data Research and Development Program" in 2012, alongside other global initiatives, underscores the importance of big data in transforming healthcare services. The Chinese government has also recognized healthcare big data as a critical national strategic resource (General Office of the State Council, 2016), promoting its integration into healthcare to enhance service quality and efficiency.

By mining and analyzing large amounts of medical data, healthcare organizations can more accurately understand patients' needs and health conditions. For example, a medical institution in the United States used data mining technology to analyze patients' consultation records, health data, and medical expenditures. This analysis revealed differences in consultation preferences and medical needs among different age groups, leading to a more targeted service strategy and significant improvements in service delivery.

Similarly, a European healthcare company utilized data mining to analyze patients' health behaviors and consumption habits. This approach provided strong data support for the positioning and promotion of healthcare services, making them more aligned with patients' needs and improving market competitiveness.

In summary, the application of data mining technology in healthcare not only supports the personalization and differentiation of medical services but also provides a scientific basis for the strategic planning of medical institutions. This facilitates the efficient implementation and high-quality delivery of healthcare services. Future research can further explore the mechanisms by which different factors influence healthcare demand and make targeted policy recommendations to promote continuous optimization and innovation in the healthcare service system.

1) Analysis of Previous Research

The aforementioned studies highlight the substantial benefits of integrating data mining technology into healthcare systems (Li & Cheng, 2012). These benefits include enhanced understanding of patient needs, improved service delivery, and more effective healthcare marketing strategies (Zhou, Li, Zhang et al., 2023).

However, these studies also exhibit several limitations:

Scope of Data: The data used in many studies is often limited to specific regions or populations, which may not fully capture the diverse healthcare needs of different demographic groups.

Data Quality and Integration: Inconsistent data quality and integration issues can lead to inaccuracies in analysis and hinder the effectiveness of data mining applications.

Privacy Concerns: There is a growing concern about the privacy and security of medical data, which can affect patient trust and willingness to share their information.

Implementation Challenges: While data mining offers valuable insights, integrating these insights into existing healthcare systems and workflows can be complex and resource-intensive.

Longitudinal Data Analysis: Many studies focus on cross-sectional data, lacking a longitudinal perspective that can provide deeper insights into the changes and trends in healthcare needs over time (Lunch, 2008).

2) Identification of Problems

These limitations point to several key issues that need to be addressed to maximize the potential of data mining in healthcare:

Comprehensive Data Collection: There is a need for more comprehensive data collection that includes diverse populations and long-term health records.

Data Standardization: Developing standardized protocols for data collection, storage, and integration can improve data quality and facilitate more accurate analyses.

Enhanced Privacy Measures: Strengthening privacy and security measures can help build patient trust and encourage broader participation in data sharing initiatives.

Efficient Implementation Strategies: Creating effective strategies for integrating data mining insights into clinical practice and healthcare management can enhance the practical utility of these technologies.

Longitudinal Studies: Conducting longitudinal studies can provide more detailed insights into the evolving healthcare needs and preferences of different populations.

3) Conclusion

In summary, the application of data mining technology in healthcare not only supports the personalization and differentiation of medical services but also provides a scientific basis for the strategic planning of medical institutions. This facilitates the efficient implementation and high-quality delivery of healthcare services. Future research can further explore the mechanisms by which different factors influence healthcare demand and make targeted policy recommendations to promote continuous optimization and innovation in the healthcare service system. Addressing the identified limitations and challenges will be crucial in realizing the full potential of data mining in transforming healthcare services.

2. Schematic Design

First, we draw a flow chart (see Figure 1).

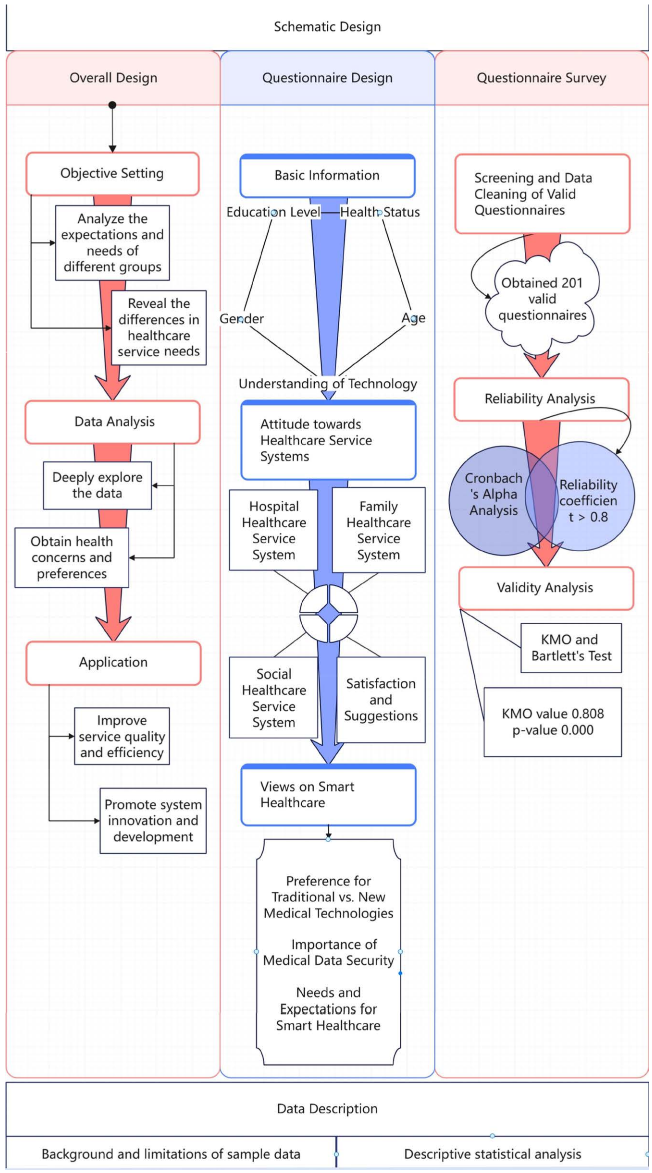

Figure 1. Framework of schematic design.

2.1. Overall Design

The goal of this study is to delve into the expectations and differentiated needs of various groups for the use of big data healthcare, thereby revealing the unique differences in healthcare service needs of different groups. Through detailed data analysis, we aim to discover the deep information behind the data and gain insights into the health concerns and preferences of different groups. This study will provide solid data support to improve the quality and efficiency of healthcare services, bring healthcare services closer to the actual needs of patients, and thus promote the innovation and development of the healthcare system.

2.2. Questionnaire Design

Through a well-designed questionnaire, this study aims to gain a comprehensive understanding of citizens' expectations for the use of big data healthcare and their individualized needs, and based on this, to provide targeted advice and support through the results of data analysis. The questionnaire is divided into three main sections:

1) Basic information

This includes gender, age, education level, health status, level of scientific and technological understanding, etc. This information helps to gain insight into the basic situation of the respondents (Batarseh & Abdel, 2016).

2) Attitudes toward different types of healthcare delivery systems

The medical service system is divided into three categories, namely, hospital, home, and social medical service system, and citizens are surveyed on their usage, satisfaction, and suggestions respectively. We also divided each service system into specific categories, explored the items with the highest level of satisfaction in that kind of healthcare service system, and then conducted regression analysis with the basic information of the surveyed people to get the user profile of the items in that category (Zhou, Tang, & Zhao, 2015).

3) Views on Smart Healthcare

The first thing that was investigated was the values and orientations, whether they preferred traditional healthcare or were more willing to try new healthcare technologies; whether they valued the security of their healthcare data, and so on, and this information helped to understand their possible rejection and acceptance factors. The second thing investigated is the needs and expectations for smart healthcare, including the problems they want these services to solve and their views on the future of smart healthcare.

2.3. Questionnaire Survey

After the steps of careful screening of valid questionnaires and data cleaning, we obtained a rich sample of data that included a total of 201 questionnaires. Subsequently, we conducted in-depth reliability and validity analyses on these data to ensure the reliability and validity of the data (Zhou, Li, & Zhang, 2023).

1) Reliability analysis

Reliability, also known as reliability, is the degree of consistency of the results obtained when repeated measurements are made on the same object. In this study, we used Cronbach's alpha as an indicator for reliability analysis to assess the internal consistency and stability of the survey instrument (see **Table 1**). After exhaustive data processing and analysis, we obtained the results of the reliability analysis as shown below: (see **Table 1**).

In reliability analysis, the coefficients are usually interpreted in different ranges. When the reliability coefficient is between 0.6 and 0.7, it means that the reliability of the questionnaire is acceptable; if it is between 0.7 and 0.8, it means that the questionnaire's reliability is relatively good; and when the coefficient is

Table 1. Reliability checklist.

dimension (math.)	Cronbach Alpha	Item count (of a consignment etc.)
In-hospital healthcare delivery system	0.883	5
Family Medical Service System	0.847	5
Social health service system	0.910	6
Scale as a whole	0.844	16

more than 0.8, it means that the quality of the data's reliability is very high. After analyzing the results, we got the result that the credibility coefficient of the questionnaire is more than 0.8, which means that the quality of the data's credibility is very high, so we can continue to carry out further statistical analysis.

2) Validity analysis

Validity refers to the accuracy of the measurement tool, that is, the degree to which the measurement results can reflect the characteristics to be measured, the higher the validity the more able to achieve the purpose of the questionnaire test. In this paper, exploratory factor analysis was conducted using KMO and Bartlett's test (see **Table 2**), the closer the KMO value is to 1, the better the correlation between the variables, and the better the effect of the factor analysis; the significance of Bartlett's test of sphericity with a P-value of significance < 0.05 does not obey the test of sphericity, which indicates that there is a strong correlation between the variables. The results were obtained as **Table 2** Shown.

Table 2. KMO and Bartlett's test table.

KMO and Bartlett's test		
KMO Number of Sampling Suitability Measure.		0.808
Bartlett's test of sphericity	approximate chi-square (math.)	2294.099
	(number of) degrees of freedom (physics)	120
	significance	0.000

From **Table 2**, the KMO value of the scale data is 0.808, which is greater than 0.5, while its significance level p-value is 0.000, which is less than 0.05. This indicates that the selected sample data is valid and there is enough correlation to achieve the expected results.

Next, the factor rotation was performed using the maximum variance method to obtain the rotated component matrix as **Table 3** shows.

From the rotated component matrix table, the three factors extracted are exactly the in-hospital health care delivery system, home health care delivery system, and social health care delivery system set in the questionnaire, i.e., it proves that the questionnaire scale is set reasonably.

Table 3. Table of rotated component matrices.

	ingredient		
	1	2	3
16. Intelligent security monitoring	0.877		
16. Health education and preventive measures	0.868		
16. Intelligent health detection	0.797		
16. Health data management	0.747		
16. Telemedicine services	0.745		
16. Intelligent medication management	0.726		
9. Clinical decision-making systems		0.866	
9. Remote consultation		0.854	
9. Automatic alarm		0.845	
9. Remote visitation		0.817	
9. Intelligent prescriptions		0.726	
12. Telemedicine			0.803
12. Health care programs			0.775
12. Disease management			0.756
12. Emergency relief			0.746
12. Health counseling			0.694
Extraction method: principal component analysis. Rotation method: Kaiser normalized maximum variance method.			
a. The rotation has converged after 5 iterations.			

3. Data Description

3.1. Data Sources

In this paper, a total of two hundred and one data were collected through the questionnaire survey. According to the background of the collected data questionnaire, the questionnaire fillers' IP addresses are concentrated in Tianjin, and the average age is about 28 - 30 years old. The limitation of age leads to our questionnaire collection problem is not comprehensive.

3.2. Variant

We then defined 10 variables based on the data collected (see Table 4).

Table 4. Variable indicator system

X_1	(a person's) age
X_2	educational attainment
X_3	health situation

Continued

X_4	Health problems faced
X_5	Proficiency in the use of electronics
X_6	Level of smart healthcare development in the region
X_7	Whether or not you believe in smart medical diagnostics
X_8	Level of importance attached to personal privacy
X_9	Is there a concern that smart healthcare platforms are invading personal privacy
X_{10}	Positive outlook for smart healthcare platforms

1) Note: In this paper, we set the variable X_i (i = 1, 2 ··· 10) variables, and all the options in the questionnaire are transformed into calculable indicator values, to facilitate the calculation in the construction of the model later.

2) Relationship with target variables: the above 10 variables are the basic characteristics of the questionnaire fillers and have a certain causal relationship with the target variables, which will be one of the reasons that constitute the fillers' choice of target variables. To facilitate the subsequent calculation, we assume that each variable has the same weight.

Descriptive Analysis of Target Variables

We then conducted a descriptive analysis of the target variables (see Table 5).

Table 5. Descriptive statistics of target variables.

descriptive statistics					
	N	minimum value	maximum values	average value	(statistics) standard deviation
In-hospital healthcare delivery system					
teleconsultation	201	3	5	4.40	0.708
automatic warning	201	3	5	4.07	0.689
Clinical Decision System	201	3	5	4.55	0.564
Smart Prescriptions	201	3	5	4.24	0.629
Family Medical Service System					
Disease management	201	3	5	4.45	0.631
bailout	201	3	5	4.67	0.542
Health Consultation	201	3	5	4.25	0.684
Health Care Programs	201	3	5	4.21	0.647

Continued

Social health service system					
Intelligent Health Detection	201	3	5	3.98	0.624
Health Data Management	201	3	5	3.93	0.591
Intelligent Medication Management	201	3	5	4.08	0.740
Intelligent Security Monitoring	201	3	5	4.17	0.672
Health education and preventive measures	201	3	5	4.16	0.607

Table 5 is our descriptive analysis based on the satisfaction ratings of the three aspects of smart healthcare systems in the questionnaire, which resulted in the smart healthcare services with the highest ratings among the three smart healthcare service systems of hospital, home, and society, respectively, which were used as the target variables, Y_1 (Clinical Decision Making System), Y_2 (Emergency Rescue), and Y_3 (Intelligent Safety Monitoring).

3.3. Descriptive Analysis of Variables

Combined with descriptive statistical graphs of variable indicators (see Table 6)

Table 6. Descriptive statistics for variables.

	N	minimum value	maximum values	average value	(statistics) standard deviation
X_1	201	1	5	2.85	1.145
X_2	201	1	7	4.19	1.666
X_3	201	2	2	2.00	0.000
X_4	201	1	3	2.24	0.613
X_5	201	1	3	2.41	0.674
X_6	201	1	3	1.40	0.610
X_7	201	1	2	1.40	0.492
X_8	201	1	2	1.16	0.367
X_9	201	1	2	1.31	0.463
X_{10}	201	1	3	1.93	0.894

Since this paper is converting the options in the questionnaire into index values that can be calculated, the minimum and maximum values in descriptive statistics, whether in variables or target variables, do not have much significance; rather, the average value is more indicative of the relationship between the basic characteristics of the filler and the target variable.

4. Model Building

4.1. Model Variable

Independent Variables (X_1 - X_{10})

X_1: Age Range: 1 to 5

Description: Represents the age group of the respondents. Categories are structured to reflect a progression from younger to older age groups, with higher values indicating older respondents.

Average Value: 2.85

X_2: Educational Attainment Range: 1 to 7

Description: Measures the highest level of education completed by the respondents. Higher values correspond to higher levels of education.

Average Value: 4.19

X_3: Health Situation Range: 2

Description: Assesses the general health status of the respondents. Since all responses are the same, this variable may indicate a homogenous health status among the participants.

Average Value: 2.00

X_4: Health Problems Faced Range: 1 to 3

Description: Indicates the severity of health issues faced by the respondents. Higher values denote more severe health problems.

Average Value: 2.24

X_5: Proficiency in the Use of Electronics Range: 1 to 3

Description: Evaluates how skilled the respondents are in using electronic devices, with higher values indicating greater proficiency.

Average Value: 2.41

X_6: Level of Smart Healthcare Development in the Region Range: 1 to 3

Description: Reflects the development level of smart healthcare infrastructure in the respondent's region. Higher values suggest better development.

Average Value: 1.40

X_7: Belief in Smart Medical Diagnostics Range: 1 to 2

Description: Measures whether respondents trust smart medical diagnostic systems. Higher values indicate stronger belief.

Average Value: 1.40

X_8: Importance Attached to Personal Privacy Range: 1 to 2

Description: Assesses how much importance respondents place on personal privacy. Higher values reflect greater importance.

Average Value: 1.16

X_9: Concern about Privacy Invasion by Smart Healthcare Platforms

Range: 1 to 2

Description: Evaluates the level of concern respondents have about smart healthcare platforms invading their privacy. Higher values denote more concern.

Average Value: 1.31

X_{10}: Positive Outlook for Smart Healthcare Platforms Range: 1 to 3

Description: Indicates the optimism of respondents regarding the future of smart healthcare platforms. Higher values represent a more positive outlook.

Average Value: 1.93

Target Variables (Y_1 - Y_3)

Y_1: Clinical Decision Making System Average Value: 4.55

Standard Deviation: 0.564

Description: Represents the satisfaction rating for the clinical decision-making systems in hospitals. Higher values indicate greater satisfaction.

Y_2: Emergency Rescue Average Value: 4.67

Standard Deviation: 0.542

Description: Measures the satisfaction rating for emergency rescue services in the family medical service system. Higher values denote higher satisfaction.

Y_3: Intelligent Safety Monitoring Average Value: 4.17

Standard Deviation: 0.672

Description: Reflects the satisfaction rating for intelligent safety monitoring services in the social health service system. Higher values suggest greater satisfaction.

These variables form the basis for building and analyzing the model, which aims to predict the satisfaction levels of different aspects of smart healthcare systems based on the demographic and personal characteristics of the respondents.

4.2. Model Detection

In this paper, after trying several machine learning classification algorithms, we decided to choose the Random Forest Regression algorithm for model construction. The following is a test of the accuracy of the model (see Table 7).

Table 7. Random forest results.

	MSE	RMSE	MAE	MAPE	R^2
training set	0.043	0.208	0.175	3.874	0.826
test set	0.587	0.766	0.657	14.058	−0.608

In Table 7, the prediction evaluation metrics of the cross-validation set, the training set, and the test set are demonstrated to measure the prediction effect of the random forest through quantitative metrics. Among them, the evaluation metrics of the cross-validation set can continuously adjust the hyperparameters to obtain a reliable and stable model (Khairo, 2014).

From Table 7, the training set MSE tends to 0, RMSE is larger but also tends to 0, and R^2 tends to 1, which can indicate that the accuracy of the model is higher, and the model predicts the results better, so this paper can utilize the model to calculate to get the prediction results, to portray the user portrait (Zhou, Leng, & Liu, 2022).

4.3. Analysis of Model Prediction Results

Since this paper portrays user profiles of smart service healthcare platforms for hospitals, homes, and societies respectively, the predictions of the following models are also divided into three parts.

4.3.1. Intelligent Medical Service Platforms within Hospitals

1) Characteristic importance ratio (see Figure 2)

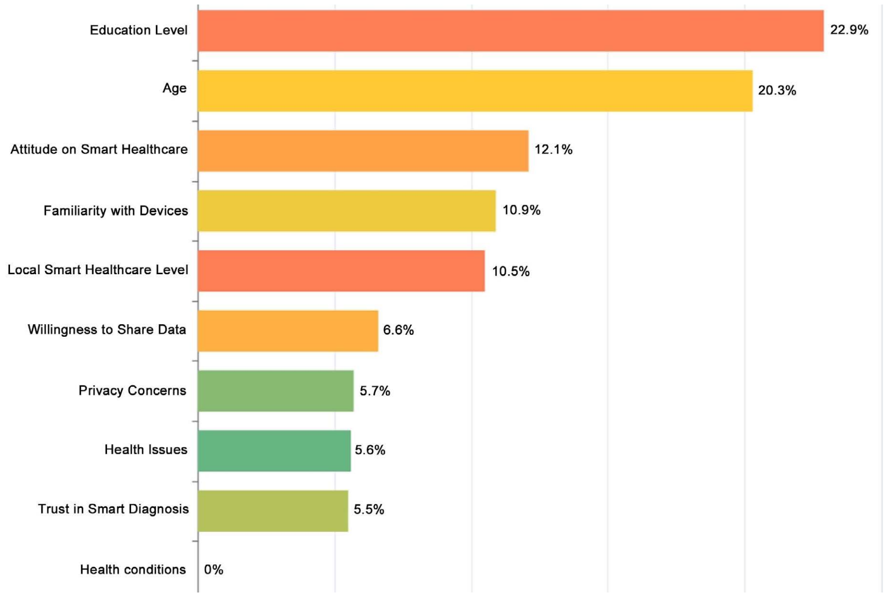

Figure 2. Proportion of importance of healthcare delivery system characteristics within hospitals.

2) Model prediction results

a) Forecast results (see Figure 3)

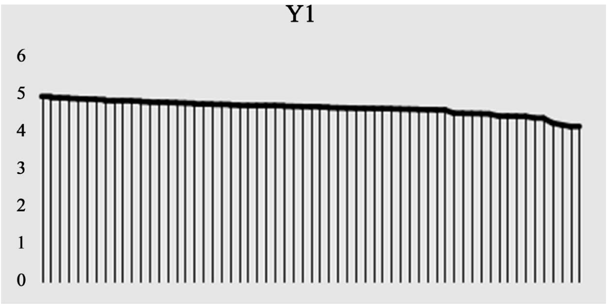

Figure 3. Results of the model prediction of the healthcare delivery system within the hospital.

According to the above figures, among the different combinations of basic features, there is a group of combinations that can achieve the highest satisfaction score of 4.98, that is to say, the questionnaire fillers with this group of combinations of basic features will give the highest satisfaction score to Y_1, which is the group of basic features that we are looking for to portray the user's profile.

b) User profile (see **Table 8**)

Table 8. Construction of user profiles for in-hospital healthcare delivery systems.

X_1	Age: 41 to 60
X_2	Level of education: Higher education
X_3	Health: certain health problems
X_4	Health problems faced: long-term
X_5	Proficient in the use of electronic devices: Yes, elders not
X_6	Level of development of smart healthcare in the region: Adequate
X_7	Do you believe in smart medical diagnosis: No
X_8	Level of importance of personal privacy: importance
X_9	Whether you are worried that the smart healthcare platform violates personal privacy: Yes
X_{10}	How optimistic is the outlook for smart healthcare platforms: Very optimistic

According to the above table, we can portray the image of users who are more willing to choose a smart medical system in hospital healthcare: middle-aged people, and elders at home who have certain long-term health problems and need to go to and from the hospital for many times, and at the same time, they have received higher education, can use smart devices more skillfully and attach more importance to their privacy, and the level of development of smart medical care is more fully developed in the region where they are located, and they have received smart medical care to a certain extent. They have received smart healthcare services to a certain extent, and are optimistic about the smart healthcare platform (Yu et al., 2024).

4.3.2. Family Intelligent Medical Service Platforms

1) Characteristic importance ratio (see **Figure 4**)

2) Model prediction results

a) Forecast results (see **Figure 5**)

According to the above figures, among the different combinations of basic features, there is a group of combinations that can achieve the highest satisfaction score of 4.91, that is to say, the questionnaire fillers with this group of combinations of basic features will give the highest satisfaction score to Y_2, and this is the group of basic features that we are looking for to portray the user's profile (see **Table 9**).

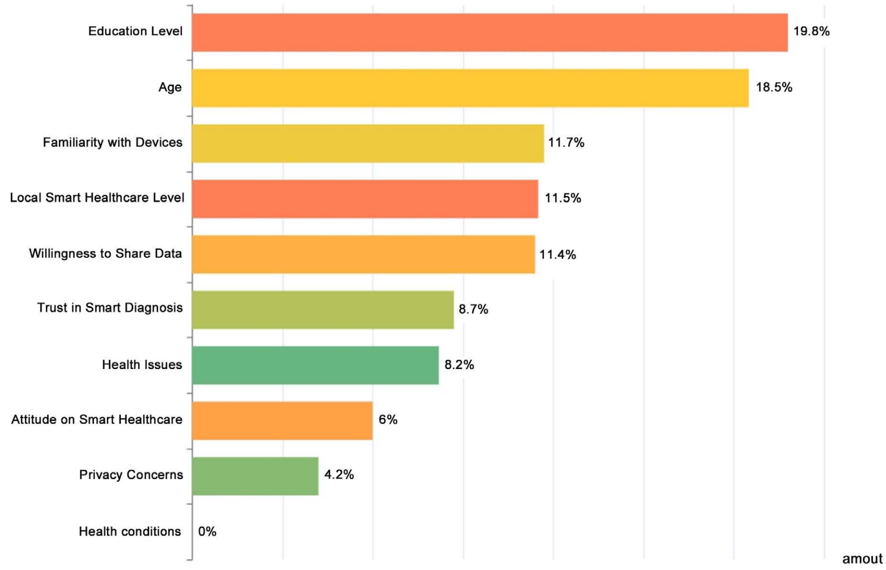

Figure 4. Proportion of importance of characteristics of the home health care delivery system.

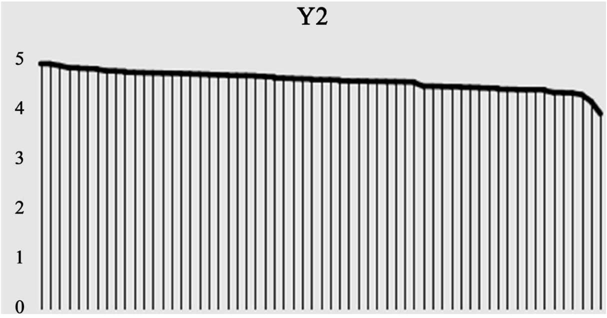

Figure 5. Results of the FMS model predictions.

b) user profile

Table 9. Construction of user profiles for home healthcare service systems.

X_1	Age: 31 - 40
X_2	Level of education: High school education
X_3	Health: certain health problems
X_4	Health problems faced: long-term
X_5	Proficiency in the use of electronics: No
X_6	Level of development of smart healthcare in the region: Adequate
X_7	Do you believe in smart medical diagnosis: No
X_8	Level of importance of personal privacy: Importance
X_9	Whether you are worried that the smart healthcare platform violates personal privacy: Yes
X_{10}	How optimistic is the outlook for smart healthcare platforms: Very optimistic

According to the above table, we can portray the portrait of users who are more willing to choose intelligent medical system in home healthcare: middle-aged people, with a medium level of education, suffering from certain long-term health problems themselves or their elders at home, and unable to use electronic products skillfully, still in the skeptical stage of intelligent medical diagnosis, but due to the high level of healthcare development in their regions, they are willing to try intelligent medical diagnosis for However, due to the high level of medical development in their region, they are willing to try smart medical diagnosis for treatment, and at the same time emphasize their privacy (Cui et al., 2023).

4.3.3. Social Intelligence Medical Service Platforms

1. Characteristic importance ratio (see Figure 6)

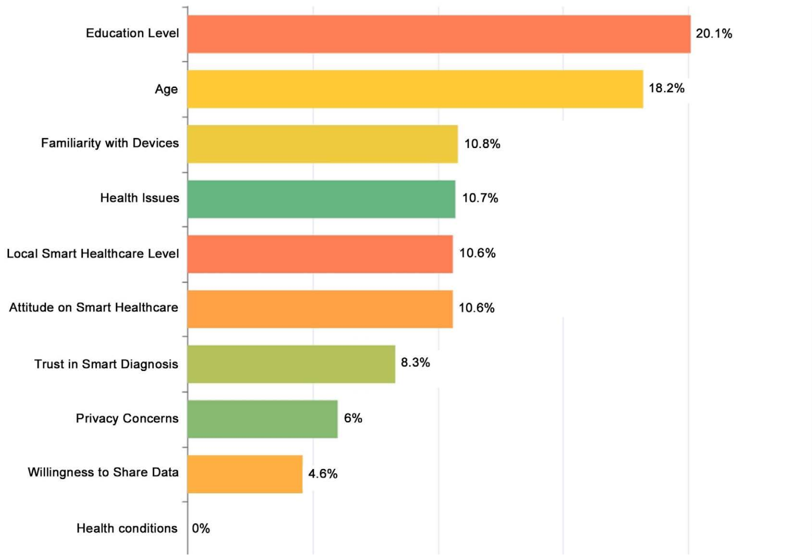

Figure 6. Proportion of importance of social health care system characteristics.

2) Model prediction results

a) Forecast results (see Figure 7)

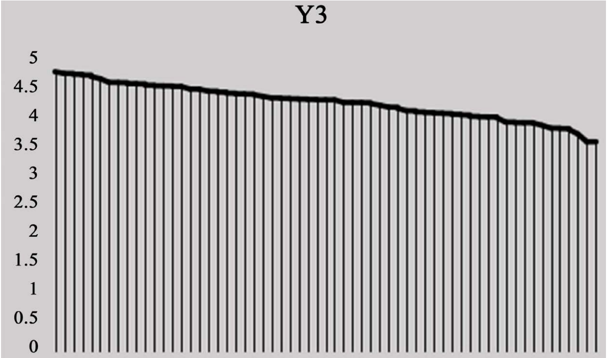

Figure 7. Results of the model predictions for the social health care delivery system.

According to the above figures, among the different combinations of basic features, there is a group of combinations that can achieve the highest satisfaction score of 4.76, that is to say, the questionnaire fillers with this group of combinations of basic features will give the highest satisfaction score to Y_3, and this is the group of basic features that we are looking for to portray the user's profile (see Table 10).

b) User profile

Table 10. Social medical service system user profile construction.

X_1	Age: 31 - 40
X_2	Level of education: Higher education
X_3	Health: certain health problems
X_4	Health problems faced: long-term
X_5	Proficiency in the use of electronics: Yes
X_6	Level of development of smart healthcare in the region: Adequate
X_7	Do you believe in smart medical diagnosis: Yes
X_8	Level of importance of personal privacy: importance
X_9	Whether you are worried that the smart healthcare platform violates personal privacy: Yes
X_{10}	How optimistic is the outlook for smart healthcare platforms: Very optimistic

According to the above table, we can portray the portrait of users who are more willing to choose smart healthcare systems in terms of social healthcare: middle-aged people, higher education, certain long-term health problems for themselves and their family members, proficient in electronic products, living in areas where the level of healthcare is more adequately developed, willing to trust the diagnosis of smart healthcare platforms, and positively optimistic about the development of smart healthcare platforms, as well as being more concerned about their privacy...

5. Summarizing the Outlook

The text collects the demand data of Tianjin residents for different healthcare service systems by means of questionnaire survey (Xiang, Zhang, & Zhao, 2021). Through the results of the questionnaire survey, the items that residents are most interested in the in-hospital medical service system, the family medical service system and the social medical service system are derived respectively, and then after trying various machine learning classification algorithms, the random forest regression algorithm is chosen for model construction, and the user profiles of the three types of medical service systems are finally obtained.

This study also has some limitations and shortcomings. First, the small sample size of the collection and the possible limitations on the richness and quality of the data sources may lead to biased results. Future research could consider integrating data from more sources to improve the accuracy and reliability of the analysis. Second, the selection and optimization of data mining algorithms is also an important research direction (Zhou, Xu, Ren et al., 2014). Currently, although we have used a variety of algorithms for analysis, there is still room for further improvement. Future research will explore more advanced and suitable algorithms for the characteristics of medical data to improve the efficiency and accuracy of analysis.

Looking ahead, the potential of data mining in the analysis of differentiated demand for healthcare services will be further realized with the continuous advancement of healthcare technology and the wide application of big data technology (Khairo, 2014). We can foresee that future research will pay more attention to the real-time and dynamic nature of data to better capture the changing trends of healthcare service demand (Kum, Ahalt, & Carsey, 2011). At the same time, with the continuous development of artificial intelligence technology, the combination of data mining with machine learning, deep learning, and other technologies will also bring more possibilities for differentiated demand analysis of medical services.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Chi Zhou, for their invaluable guidance, unwavering support, and insightful feedback throughout the entire process of researching and writing this paper. Their expertise, encouragement, and patience have been instrumental in shaping the direction of this work and refining its content.

I am also indebted to my classmates, whose camaraderie, encouragement, and exchange of ideas have enriched my academic journey. Their constructive criticism and diverse perspectives have contributed significantly to the development and improvement of this paper.

Additionally, I extend my appreciation to Kai Gao, Yihao Yin, Xin Li, and Jing Yang for their assistance with data collection and analysis, as well as their willingness to engage in thoughtful discussions that have deepened my understanding of the subject matter.

Finally, I would like to thank my family and friends for their unwavering support, understanding, and encouragement throughout this endeavor. Their love and encouragement have been a constant source of motivation and inspiration.

This research would not have been possible without the generous support and contributions of all those mentioned above, and for that, I am sincerely grateful.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Batarseh, F. A., & Abdel Latif, E. (2016). Assessing the Quality of Service Using Big Data Analytics: With Application to Healthcare. *Big Data Research, 4,* 13-24. https://doi.org/10.1016/j.bdr.2015.10.001

Cui, X. et al. (2023). Interaction between Manufacturer's Recycling Strategy and e-Commerce Platform's Extended Warranty Service. *Journal of Cleaner Production, 399,* Article ID: 136659. https://doi.org/10.1016/j.jclepro.2023.136659

General Office of the State Council (2016). *Guiding Opinions on Promoting and Regulating the Development of Healthcare Big Data Applications State Office Development.* http://www.gov.en/zhengce/content/2016-06/24/content_5085091.htm

Khairo, M. (2014). Mobile Data Mining-Based Services on the Base of Mobile Device Management (MDM) System. *Journal of Signal and Information Processing, 5,* 89-96. https://doi.org/10.4236/jsip.2014.53011

Kum, H. C., Ahalt, S., & Carsey, T. M. (2011). Dealing with Data: Governments Records. *Science, 332,* 1263. https://doi.org/10.1126/science.332.6035.1263-a

Li, G. J., & Cheng, X. Q. (2012). Big Data Research: A Major Strategic Field for Future Scientific and Technological and Economic and Social Development—Research Status and Scientific Thinking of Big Data. *Proceedings of the Chinese Academy of Sciences, 27,* 647-657. (In Chinese)

Lunch, C. (2008). Big Data: How Do Your Data Grow? *Nature, 455,* 28-29. https://doi.org/10.1038/455028a

National Health and Family Planning Commission (2017). *Notice on the Issuance of the "13th Five-Year Plan" for the Development of National Population Health Informatization.* http://www.nhfpc.gov.cn/guihuaxxs/s10741/201702/ef9ba6fbe2ef46a49c333de32275074f.shtml

Xiang, K. F., Zhang, C. F., & Zhao, S. L. (2021). Clinical Analysis of 76 Cases of Differentiated Thyroid Isthmic Carcinoma after Operation. *Open Access Library Journal, 8,* 1-8. https://doi.org/10.4236/oalib.1107350

Yu, J. et al. (2024). Is It Always Advantageous to Establish Self-Built Logistics for Online Platforms in a Competitive Retailing Setting? *IEEE Transactions on Engineering Management, 71,* 1726-1743. https://doi.org/10.1109/TEM.2023.3337241

Zhou, C., Leng, M., Liu, Z. et al. (2022). The Impact of Recommender Systems and Pricing Strategies on Brand Competition and Consumer Search. *Electronic Commerce Research and Applications, 53,* Article ID: 101144. https://doi.org/10.1016/j.elerap.2022.101144

Zhou, C., Li, H., Zhang, L. et al. (2023). Optimal Recommendation Strategies for AI-Powered e-Commerce Platforms: A Study of Duopoly Manufacturers and Market Competition. *Journal of Theoretical and Applied Electronic Commerce Research, 18,* 1086-1106. https://doi.org/10.3390/jtaer18020055

Zhou, C., Tang, W., & Zhao, R. (2015). Optimal Consumer Search with Prospect Utility in Hybrid Uncertain Environment. *Journal of Uncertainty Analysis and Applications, 3,* Article No. 6. https://doi.org/10.1186/s40467-015-0030-z

Zhou, C., Xu, Y., Ren, Y. et al. (2024). Strategic Adoption of the Recommender System under Online Retailer Competition and Consumer Search. *Electronic Commerce Research and Applications, 64,* Article ID: 101376. https://doi.org/10.1016/j.elerap.2024.101376