A Study on the Location Planning and Layout Prediction of Elderly Care Facilities in the Central Urban Area of Lanzhou City Based on A Machine Learning Model

Yinxia Tian and Tianpeng Wang

School of Architecture and Urban Planning, Lanzhou Jiaotong University, Lanzhou, Gansu, China

mail: 3489229804@qq.com; * Correspondence:wangtp@mail.lzjtu.cn

Abstract. Based on POI data and "seven universal" population and other data, using the decision tree model based on ID3 algorithm, according to the basic research unit of 500 meters × 500 meters, 4,466 research units are divided into the central urban area of Lanzhou City for machine learning simulation and training, and taking into account the characteristics of the distribution of the current urban senior care facilities, we get 961 suitable site distribution units. 961 suitable site distribution units, on the basis of which quantitative simulation predictions are made for the site layout of senior care facilities in the study area. The results show that in the future, senior care facilities can be located in the following key research units: the southern part of Hekou Town, the central part of Dachuan Town, and the northern part of Xincheng Town in Xigu District; the central part of Xiguoyuan Town, the central part of Huangyu Town, and the central part of Bali Town in Qilihe District; the central part of Shajingyi Street and the northwestern part of Anningbao Street in Anning District; and the western part of Volunping Street, the western part of Yancanglu Street, the northern part of Caochang Street, and the northern part of Jingyuan Street in Chengguan District.

Keywords: poi data; machine learning model; decision tree; senior living facilities; planning layout prediction

1 Introduction

As the process of population ageing in China accelerates, the pressure on social security and public services increases, and population development faces major risks and challenges.[1] The population development is facing major risks and challenges. According to the United Nations standards, when the number of people aged 60 and above exceeds 10 percent of the country or region, or when the number of people aged 65 and above exceeds 7 percent of the country or region, it indicates that the population of the country or region in general is in a state of aging.[2] . Against the backdrop of an aging population, the issue of old age has received increasing attention.[1] . Northwest China is an underdeveloped region in the northwest of China, and

M. F. b. Sedon et al. (eds.), *Proceedings of the 2024 3rd International Conference on Social Sciences and Humanities and Arts (SSHA 2024)*, Advances in Social Science, Education and Humanities Research 851,
https://doi.org/10.2991/978-2-38476-259-0_105

at present, Northwest China is facing many problems such as serious population aging, insufficient supply of elderly care services, and poor service quality[3] . According to the data released by Lanzhou Municipal Bureau of Statistics on February 15, 2023, Lanzhou City has a total of 4,415,300 permanent residents, of which there are 815,000 elderly people aged 60 years and above, accounting for 18.46% of the city's permanent population, and 561,200 elderly people aged 65 years and above, accounting for 12.71% of the city's permanent population. In recent years, with the continuous development of spatial data mining technology, more and more complex problems have been effectively solved[4] . Therefore, the research object of this article is the central urban area of Lanzhou City, based on POI (point of interest) data, using machine learning models and other methods, combined with ArcGIS and other software, to carry out simulation research on the spatial layout of urban elderly services, with a view to improving the global, scientific and objective spatial layout, and to provide a basis for the governmental departments to formulate the relevant policies.

On the whole, for the spatial layout of senior care facilities, the current research is mainly based on qualitative research and traditional algorithms, which is influenced by subjectivity and policy and lacks scientificity. In view of this, this article will combine algorithms such as interest points and machine learning to establish a spatial siting model for senior care facilities suitable for China's national conditions, to improve the global, scientific and objective nature of spatial location, and to carry out empirical research in combination with the actual situation of Lanzhou's downtown area, with a view to enriching and developing the spatial layout theories and methods for senior care facilities.

2 Study area and research methodology

2.1 Study area

Lanzhou is the capital of Gansu Province and a major town in Northwest China. In view of the geomorphological characteristics of "two mountains and one river" in Lanzhou city, the settlements are mainly concentrated in the flat areas along the Yellow River, and the four districts and counties of Chengguan, Anning, Qilihe and Xigu, i.e., the central urban area of Lanzhou City, are densely populated. According to the "Gansu Province 7th Population Census Bulletin" , the resident population of Lanzhou City is 4,359,446, of which 16.56% are aged 60 or above and 11.7% are aged 65 or above. It can be seen that the number of people older than or equal to 65 years old exceeds the international standard by 7 %, which is a sign that Lanzhou City has entered an aging society. By the end of 2020, there are 43 institutional nursing institutions in Lanzhou City, totaling 14,290 beds; 857 community nursing institutions, with 25,710 beds, totaling 40,000, with an average number of 6 beds per 1,000 old people, and there is a shortage of nursing facilities. Therefore, this study intends to scientifically and reasonably select the location of the newly built nursing facilities and optimize their allocation. The study area of this thesis is 4 districts and 62 streets (towns) in the central city of Lanzhou.

2.2 Research methodology

The research framework of this paper is shown in Fig. 1. Firstly, based on the spatial distribution characteristics of the existing senior living facilities in the central city of Lanzhou obtained from the analysis of POI data, and with reference to the service radius of the general senior living facilities, we take the 500m× 500m analytical grid as a basic research unit, and adopt the machine learning model to optimize the living environment for the elderly in the central city of Lanzhou.[5-6-7] to optimize the living environment of the elderly in Lanzhou city center, and obtain the optimization decision-making model applicable to the prediction of the layout of the elderly facilities in Lanzhou city center. Then, the reasonableness and scientificity of the decision-making model are examined according to the layout of the existing elderly facilities, and the model is proved to be credible when the match between the existing elderly facilities and the simulation data is more than 80%. Finally, the model is applied to simulate the prediction of each research unit to determine the suitable neighborhoods for new senior care facilities.

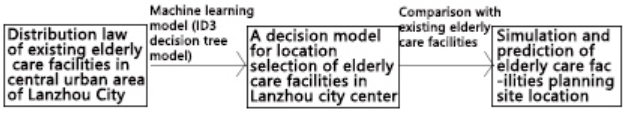

Fig. 1. Research framework diagram

The basic idea of machine learning is that the use of a certain number of training samples makes it possible for a machine to use these samples to analyze unknown data. Machine learning is about allowing a machine to learn from a large amount of data to come up with a model that is closer to the reality of the situation.

ID3 algorithm is a classification prediction algorithm, originally proposed by Prof. Roskun of the University of Sydney in 1975, and its core idea is "information entropy". The core idea is "information entropy". Information entropy and information quantity are used as metrics to categorize data inductively.[8] This article will use the ID3 algorithm to categorize the data in a generalized way. This article will be based on the ID3 algorithm, using the existing sample as an example, to construct a decision tree model.

The information entropy of D, $info(D)$, is given by the sample set $S=\{S_1,S_2,\cdots,S_n\}$ with sample number S and its set of decision attributes $D=\{D_1,D_2,\cdots,D_{\mathrm{m}}\}$ and the jth class attribute $D(ij=1,2,\cdots,m)$ with sample number $|D_j|$:

$$info(D)=-\sum_{j=1}^{m}p_j\cdot iog2\,p_j \tag{1}$$

where the probability of the jth type is p_j , which is usually estimated as the ratio of the number of samples in this category to the total number of samples. If attribute A has a set $A=\{A_1, A_2, \cdots, A_k\}$, where the value of its ith conditional attribute is $A=(i=1,2,\cdots,k)$, where the number of samples to take the value is $|A_j|$, and the number of samples of the jth type of the next A_i is $|A_{ij}|$, then the entropy of information after partitioning attribute A is:

$$info(A)=-\sum_{i=1}^{k}\frac{|A_i|}{|s|}\cdot\sum_{j=1}^{m}p_{ij}\cdot\log 2\,p_{ij} \tag{2}$$

Finally, the information gain (i.e., the difference in information entropy) is obtained before and after the segmentation of attribute A, denoted as:

$$Gain(A)=info(A)-info(D) \tag{3}$$

Decision Tree (Decision Tree) is to evaluate the project and determine its feasibility based on the probability of occurrence of different scenarios by constructing a decision tree to get the probability that the expectation of the net present value is not less than zero. It is a graphical method that directly utilizes probabilistic analysis[9] . Based on the decision tree model under the ID3 algorithm, the basic idea is to sequentially perform the information increment of each attribute before and after segmentation, and each time the largest information increment is taken as the node, that is, the most efficient attribute is taken as the classification condition, and the optimal decision tree model is finally obtained through many iterations.

3 Data sources and Data processing

3.1 POI data

The "Points of Interest" data used in this paper comes from the API (Application ProgrammingInterface) given by the developers of Gaode Map, and is obtained using the Amap_poi tool. The collected POI data is in the form of a database, including the name of the facility, functional category, latitude and longitude, and other basic information. The coordinate system of latitude and longitude established by using Gaode map was converted using ArcGIS.10.8.1 software to get the coordinates of WGS84.

3.2 Demographic data

The data of the elderly population selected in this paper is from the "seven universal" data of Lanzhou city center, subdivided into 62 streets (townships) in Lanzhou city center. Statistical data show that there are five streets in Lanzhou city center with

more than 10,000 elderly people aged 65 or older, namely, Xihu, Dunhuang Road, West Road, Pioneer Road and Fuli Road. Among these streets, there are 37 streets with more than 5,000 elderly people older than or equal to 65 years old, accounting for 59.68 % of the total number of all streets; there are 47 streets with more than 3,000 elderly people older than or equal to 65 years old, accounting for 75.81 % of the total number of all streets. The distribution of the elderly population in the central city of Lanzhou starts from the northern area and gradually spreads to other areas. (Figure 2).

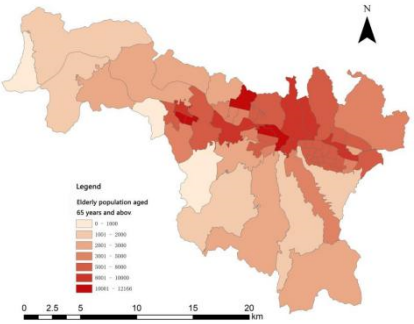

Fig. 2. Distribution of aging population in each street in Lanzhou city center

3.3 Classification of infrastructure and division of study units

First, various types of public service facilities related to senior living facilities in the city are categorized based on the relationship of senior living facilities and characterized by points of interest. Fourteen categories are included, namely: transportation facilities, science, education and culture, sports and leisure, scenic spots, living services, finance and insurance, shopping services, lodging services, food and beverage services, companies and enterprises, governmental institutions, business residences, healthcare services and senior living facilities, totaling 58,242 entries (as shown in Table 1). These data cover 4 districts and 62 streets (townships) in Lanzhou city center.

Table 1. POI data of various types of public service facilities associated with elderly care facilities

Serial number	Form	Name of facility	Number of POIs (pcs)
1	transportation facilities	Transportation services related to airports, train stations, port terminals, long-distance bus stations, subway stations, light rail stations, bus stations, shuttle bus stations, parking lots, border crossings, cabs, ferry stations, ropeway stations, loading and unloading areas, etc.	3045
2	science education culture	Science, education and cultural venues, museums, exhibition halls, convention and exhibition centers, art museums, libraries, science and technology museums, planetariums, cultural	2303

		palaces, archives, arts and culture groups, media organizations, schools, scientific research institutes, training institutions, driving schools, etc.	
3	Sports and Leisure	Sports and leisure service venues, sports venues, entertainment venues, vacation and recreation venues, leisure venues, theaters, etc.	1560
4	scenic spots	Scenic spots related, parks and squares, scenic spots, etc.	365
5	Life Services	Life service places, travel agencies, information centers, ticketing offices, post offices, logistics couriers, telecommunication business offices, offices, job markets, water business offices, electric power business offices, beauty salons, laundromats, intermediaries, sharing facilities, etc.	8553
6	financial insurance	Financial and insurance service organizations, banks, bank-related, ATMs, insurance companies, securities companies, finance companies, etc.	869
7	Accommodation services	Accommodation service related, hotels, guest houses, etc.	2431
8	Shopping Services	Shopping malls, convenience stores, home appliances and electronic stores, supermarkets, flower, bird, fish and insect markets, home building materials markets, general markets, cultural goods stores, sporting goods stores, specialty shopping streets, clothing, shoes, hats, and leather goods stores, and special trading venues, etc.	17791
9	Catering Services	Chinese restaurants, foreign restaurants, fast food restaurants, casual dining places, coffee shops, tea houses, dessert stores, etc.	10003
10	government organization	Government agencies, foreign institutions, democratic parties, social organizations, public prosecutors and law enforcement agencies, transportation vehicle management, etc.	2580
11	Company	Corporate enterprises, well-known enterprises, companies, factories, agriculture, forestry and fishery bases, etc.	4491
12	Business Residential	Business and housing related, industrial parks, buildings, residential areas, etc.	1921
13	health care	Healthcare service places, general hospitals, specialized hospitals, clinics, emergency centers, disease prevention agencies, etc.	2207
14	Retirement facilities	Nursing homes, elderly care centers, homes for the aged, senior citizen universities, senior citizen apartments, care centers, social welfare centers, etc.	123
Total	58242		

Data source: Gaode Map Developer API Interface (https://lbs.amap.com/api/webservice/download)

After that, the length of a 15-minute walk for the elderly was used as the benchmark, the service radius of ordinary elderly facilities was used as the reference, and the walking distance of the elderly was used as the reference, and the central city of Lanzhou was divided into 4,466 study units by using 500 m*500 m as a study unit. Each study unit has its own unique numerical code, which facilitates the exploration of the layout of various facilities. For example, there is a community elderly culture room in Chengguan District, Lanzhou city center, in research unit 2200, which in turn has many facilities such as State Grid Gansu Electric Power Company, Caochangjie Street Office, and Haosheng International Fitness at the same time (shown in Figure 3).

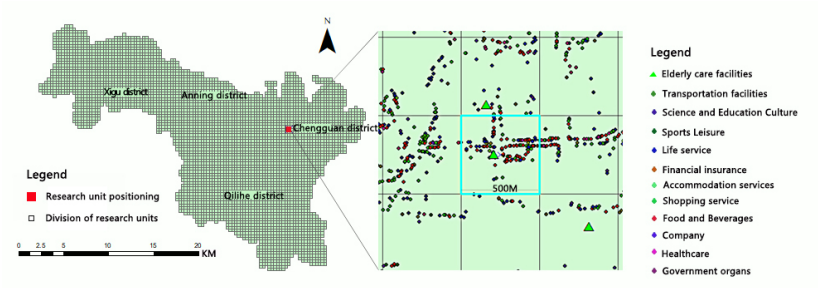

Fig. 3. Study unit division and localized schematic diagram of urban facilities in Lanzhou city center

Finally, ArcGIS and Python programming were used to realize the distribution of various types of facilities in each research unit. In the case of research unit 2200 mentioned above, there are "elderly facilities, transportation facilities, sports and leisure, living services, science, education and culture, financial and insurance services, lodging services, food and beverage services, shopping services, governmental organizations, companies and healthcare services". By analyzing the spatial distribution of each study unit and each facility, it is possible to determine whether a facility exists in each study unit.

4 Predictive analysis

First of all, the study is based on the spatial layout characteristics of existing senior care facilities, combined with the previous research.9 First, based on the spatial layout characteristics of existing senior care facilities, the study was combined with previous studies to form a decision tree model based on the ID3 algorithm, which was trained to obtain an optimization model for the layout of senior care facilities in the central city of Lanzhou. On the basis of the trained decision tree model, simulations and predictions were made for all the research units to determine whether they are suitable for the future planning and construction of new senior care facilities.

4.1 Decision Tree Modeling for Senior Living Facility Siting

First, research units and facilities were categorized and matched to determine the suitability of senior living facilities in a research unit, i.e., if there is a senior living facility in a research unit, it means that the objective environment of this research unit meets the spatial layout needs of this research unit, and the research unit was classified as a positive category, i.e., positive category. If there is no elderly care facility in a research unit, then the objective environment of the research unit does not meet the layout needs of elderly care facilities, and the research unit is categorized as a negative category, i.e., negative category. Thus, the purpose of secondary categorization of senior living facilities is achieved. Then, the other kinds of facilities in the research unit are also categorized and transformed into secondary categories respectively, and

if there exists a certain type of facility, then this type of facility is a positive category, and if there is no such type of facility, then it is a negative category.

The layout decision of senior care facilities is to continuously put forward the demand for each candidate site and eliminate the sites that do not meet the demand, and this process is very much in line with the idea of the decision tree model. The research object of this paper is the central city of Lanzhou, using ID3 algorithm, the existing elderly service facilities, the presence of elderly service facilities and other factors, through the learning of the decision tree, to get the decision tree model (Figure 4).

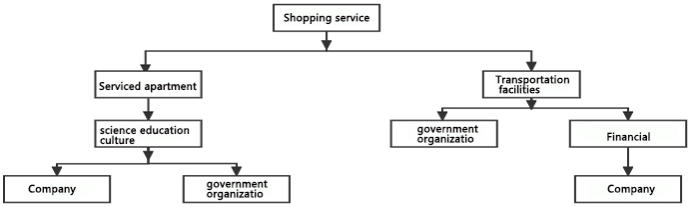

Fig. 4. Decision Tree Model for Siting of Senior Living Facilities

As can be seen from the results in Figure 4, the left and right sides of the decision tree model represent the presence or absence of such facilities in this research unit and, according to the principles of the ID3 algorithm, the higher up the decision tree model it is located, the more important it is for decision making.

The first level of the decision tree is "shopping services", which indicates that "shopping services" is the most important factor in the location of elderly facilities, and the need of "the elderly" for "shopping nearby" leads to the layout of "existing" social facilities. The need of "elderly" for "shopping nearby" leads to the "existing" layout of "social facilities".

The second tier on the left is business residential, indicating that if the senior living facility is set up in a study unit without shopping services, then priority will be given to building closer to the real estate, such as community nursing homes and home care facilities, to satisfy the siting requirements of the senior living facility. The second tier on the right, on the other hand, means that if the senior living facility is set up in a study unit with a shopping center, then there will be more transportation facilities in this study unit to better meet the travel needs of the elderly.

The left side of the third level is science, education and culture, which meets the needs of the elderly to carry out learning. On the right side of the third level are institutions and financial and insurance services, referring to the fact that there will also be financial and insurance service facilities next to the senior living facilities, such as insurance companies, and some government departments, such as community senior living facilities, will be set up in community service stations.

The fourth level and several types of service facilities not shown in Figure 4 have less influence on the decision of the location of senior living facilities, such as corporate enterprises in level 4, suggesting that some senior living facilities will be set up in areas where there are corporations, but senior living facilities may also be set up in more remote areas, such as many senior living estates and orphanages built on the

outskirts of the city. For some facilities, such as living services, sports and leisure and life services, senior living facilities are less dependent.

4.2 Projections for senior living facility site planning

Firstly, all the data were screened and sampled to ensure that the training set did not overlap with the prediction set, and among the 4466 valid research units, 100 (50 containing elderly care facilities and 50 not containing elderly care facilities) valid research units were randomly selected as the sample training set using equal positive and negative sampling rates. Taking Lanzhou central city as the research object, using the senior care facilities in each research unit as the dependent variable and the remaining 13 types of facilities as the independent variables, the decision tree model was used to train the spatial conformity model prediction for all 4,466 research units in Lanzhou central city, which in turn resulted in the 961 research units conforming to the spatial layout of the central city of Lanzhou in terms of siting, as shown in Figure 5.

The model is then validated, using the existing 123 senior living facilities as a sample test set, and the accuracy test is conducted, and the final accuracy reaches 83.3 %, exceeding the 80 % match, indicating that the final prediction results are more reasonable.

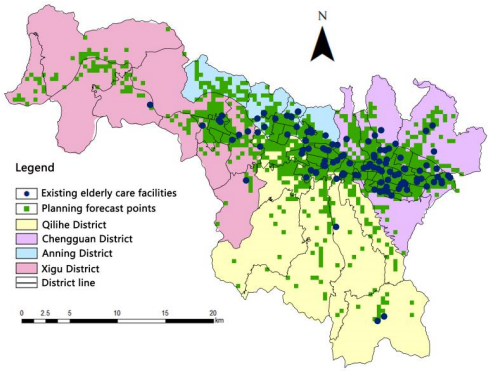

Fig. 5. Forecast results of the planning layout of senior care facilities in Lanzhou's downtown area

It can be found from the results of the previous prediction in Figure 5 that the distribution of existing senior care facilities in the central city of Lanzhou has distinctive characteristics, mainly concentrated in the central and northern parts of the city, and concentrated towards the center, with sporadic distribution in other areas. Most of the existing elderly facilities are located in the planned study units, but there are still many communities that have not yet been built in the study units that are expected to be suitable for construction, suggesting that elderly facilities can be set up in this study unit in the future, such as the southern part of Hekou Town, the central part of Dachuan Town, and the northern part of Xincheng Town in Xigu District; the central

part of Xiguoyuan Town, the central part of Huangyu Town, and the central part of Bali Town in Qilihe District; the central part of Shajingyi Street, the northwestern part of Anning Fort Street, and the northwestern part of Volong Street in Chengguan District. In the northwest of Anning District, the western part of Fulongping Street, the western part of Yancheng Road Street, the northern part of Caochang Street, and the northern part of Jingyuan Road Street in Chengguan District, there is a large amount of study unit space available for the selection of senior care facilities for the elderly to reside in.

5 Conclusion

The study takes the central city of Lanzhou as the research object, based on the population data and POI data of "seven universal", and adopts the machine learning model (decision tree model under ID3 algorithm) to simulate it, and obtains the following conclusions:

(1) The study of the central city of Lanzhou City, 4466 effective research units, the use of machine learning model to train the model, the first time to get 961 suitable deployment units, and with the existing 123 pension facilities for the accuracy of the test, the accuracy of 83.3 %, indicating that the method is feasible.

(2) At present, it is expected that there are still many communities that have not yet been built in the research units suitable for construction, suggesting that senior living facilities can be set up in the research units in the future, for example, the southern part of Hekou Town, the central part of Dachuan Town, and the northern part of Xincheng Town of Xigu District; the central part of Xiguoyuan Town, the central part of Huangyu Town, and the central part of Bali Town of Qilihe District; the central part of Shajingyi Street, and the northwestern part of Anningbao Street of Anning District; the western part of Fulongping Street, western part of Yancanglu Street, northern part of Caoba Street, and northern part of Jingyuan Road Street of Chengguan District. Street in the west, Caochang Street in the north, and Jingyuan Road Street in the north, etc. There is a large amount of study unit space available for the siting of senior living facilities for the elderly. Therefore, it is possible to provide some suggestions for the subsequent construction of senior living facilities.

(3) Point-of-interest data has become an important basis for urban planning and spatial optimization planning with its advantages of rich content, information and timeliness. This study applies the machine learning method to the planning and siting of senior living facilities, which makes the whole layout tend to global optimization, and makes the layout of senior living facilities more reasonable and more in line with the local reality, so as to reduce the designer's subjectivity, and improve the scientific and global nature of the siting.

References

1. Fengling Xiao, Du Hongru, Zhang Xiaolei. Evaluation of spatial allocation of residential neighborhoods and public service facilities under the perspective of "15-minuteliving circle"--Taking Urumqi City as an example[J]. Arid Zone Geography, 2021,44(02):574-583.
2. Xuesong Yao, Leng Hong, Wei Ye, et al. Evaluation of urban park supply based on the activity demand of the elderly--taking the main urban area of Changchun as an example [J]. Economic Geography, 2015, 35(11): 218 -224.
3. Yan B，Gao X，Lyon M. Modeling satisfaction amongst the el-derly in different Chinese urban neighborhoods［J］. Social Sci-ence & Medicine，2014，118：127 - 134.
4. Ishikawa N, Fukushige M. Dissatisfaction with dwelling envi-ronments in an aging society: an empirical analysis of the Kanto Area in Japan [J]. Review of Urban & Regional Development Studies, 2015, 27(3): 149 - 176.
5. Singelenberg J, Stolarz H, McCall M E. Integrated service ar-eas: an innovative approach to housing, services and supports for older persons ageing in place[J]. Journal of Community & Applied Social Psychology, 2014, 24(1):69 - 73.
6. Joseph A E，Skinner M W. Voluntarism as a mediator of the ex-perience of growing old in evolving rural spaces and changing rural places［J］. Journal of Rural Studies，2012，28(4)：380- 388.
7. Hongjin,Qin Gan. Research on the correlation between urban open space morphological elements and air temperature in severe cold areas based on tree model[J]. Building Science,2023,39(08):1-9+18.
8. Tao Z, Cheng Y, and Liu J. 2020. Hierarchical Two-Step Floating Catchment Area (2SFCA) Method: Measuring the Spatial Accessibility to Hierarchical Healthcare Facilities in Shenzhen, China.International Journal for Equity in Health,19 (1): 164-179.
9. Taoyu chen,AN Haiyan,Chen Jie. Analysis of factors affecting urban integration of migrant workers based on ID3 algorithm[J]. Software Engineering,2023,26(10):45-48.
10. Xiao Wang, Zhang Chaoyong, Ren Xueling, Duan Xiaoxia, Peng Hongxia, Tang Xueqin, Ding Ling, Chen Danyi, Yang Yanfang, Zrng Yilan, Long Lu. Research on drgs grouping of hand-foot-mouth disease hospitalized patients based on decision tree model[J]. Modern Preventive Medicine,2023,50(01):121-126+158.