Journal of Education and Educational Research
ISSN: 2957-9465 | Vol. 8, No. 3, 2024

Analysis of Influencing Factors of Residents' Happiness in Smart Cities based on Principal Component Analysis

Jingkun Tian

Anhui University of Finance and Economics, Bengbu Anhui, 233030, China

Abstract: This paper analyzes the factors affecting the happiness of smart city residents, uses principal component analysis method and random forest method, and uses SPSS, SPSSPRO and other software to analyze and process the data. First, it establishes the evaluation index of the happiness of smart city residents, and then evaluates the measures that can improve the happiness of residents in the construction of smart city. Finally, reasonable development suggestions for the construction of smart cities are put forward through forecasting and processing. Research background and theoretical literature. Find out the relevant theoretical basis, and finally determine the reasonable research method. The index of resident happiness is designed from the perspective of smart economic happiness, smart political happiness, smart public service happiness, smart security management happiness, and smart ecological environment happiness, and is used as the index factor of principal component analysis. The model of influencing urban residents' happiness under smart city. According to the resident data, the principal component analysis method is constructed, and then the random forest model is used to test, and finally the weight of the influencing factors of residents' happiness under the smart city is obtained.

Keywords: Smart Cities; Resident Blessedness; Principal Component Analysis; Random Forest Regression.

1. Introduction

1.1. Research Background

As a product of the new era of urban development, the Smart City aims to achieve more scientific development, efficient management, and a better living environment. Relying on advanced information and communication technologies, it significantly enhances the operational efficiency of the city and greatly optimizes and improves the quality of public services through transparent access to information, widespread and secure transmission, as well as effective and scientific processing. The construction of Smart Cities not only contributes to the formation of a low-carbon and environmentally friendly urban ecosystem, but also establishes a new type of urban form, representing the future direction of urban development. Since the concept of the Smart City emerged in 2008, it has attracted widespread attention in the international community and become a global trend in Smart City development. Nowadays, the Smart City has become a key strategy to promote global urbanization, serving as an important engine for enhancing urban governance, addressing development challenges in large cities, optimizing public service experiences, and promoting the vigorous development of the digital economy [1].

1.2. Research Status

With the continuous advancement of Smart City construction, its connection with residents' daily lives is also becoming increasingly close, attracting widespread attention from foreign researchers. Vijcendcr and other scholars have conducted in-depth research and found that the application of cutting-edge digital and information communication technologies is effective in improving the quality and effectiveness of Smart City services, significantly enhancing residents' urban living experience and satisfaction[2].

Cottrell and other scholars selected natural parks as typical cases and conducted in-depth analysis of multiple dimensions such as economic, environmental, institutional, and sociocultural sustainability. They discussed in detail the specific impacts of these sustainability factors on residents' well-being in promoting sustainable tourism development strategies[3]. Their research provides valuable references for understanding the influencing factors of residents' well-being perception in the context of sustainable development.

Bjornskov and others conducted in-depth analysis and discussion on the factors that affect happiness from institutional perspectives. They classified economic and political institutions as two types of institutional factors and empirically proved their impact on residents' happiness [4]. This helps us gain a deeper understanding of the influence of institutional factors on residents' subjective well-being.

In addition, Mccrea and others took the Queensland region as an example to explore the impact of factors such as basic supporting service facilities, sense of security, and living costs on residents' happiness in their living environment [5]. They found that these factors have significant impacts on residents' happiness, providing key considerations for policymaking in Smart Cities.

In summary, researchers have conducted multi-angle and multi-level studies on the relationship between Smart Cities and residents' happiness. These studies not only help us deeply understand the connotation and value of Smart Cities, but also bring profound insights and beneficial guidance for enhancing residents' satisfaction and promoting the sustainable development of cities. The core concept of the Smart City lies in taking human needs as the starting point and endpoint, so its essence is to promote the modernization and intelligence of urban development through reform and innovation. Therefore, in English, the "Innovative Smarter City" should be used to translate the new-type Smart City, rather than simply "New Smart City," to better reflect its inherent spirit and characteristics [6].

2. Analysis of Influencing Factors Using Principal Component Analysis

2.1. Symbol Explanation

The following is a symbol explanation in the principal component analysis, as shown in Table 1

Table 1. Symbol Explanation

Symbol	Explain
x_j	Index data
r_{ij}	Coefficient of correlation between variables
R	Correlation between indicators
λ_i	The eigenvalue of R
α_i	Eigenvector corresponding to λ_i
ϕ_i	variance contribution rate
β_i	Factor contribution rate
l_{ij}	Principal component load

2.2. Introduction to the Principal Component Analysis Model

Principal Component Analysis (PCA) is a data analysis method whose core idea lies in reducing the dimensionality of multiple indicators to extract a few comprehensive indicators. These comprehensive indicators can maximize the retention of key information in the original data, minimizing information loss while reducing the data dimensions. Through PCA, we can simplify the data, making it easier to process and analyze[8]. Given the excessive dimensionality of the current issue, this article utilizes principal component analysis to comprehensively analyze residents' happiness in smart cities, thoroughly assessing and comparing the factors that affect residents' happiness. Below are the steps of the principal component analysis model.

Step 1: First, perform data standardization. Taking into account the existence of different indicators, this article collects data for each indicator and sets it as $x_j = \left(x_{1j}, x_{2j}, x_{3j}, ..., x_{nj}\right), j = 1, 2, 3, ..., p$. To eliminate the difference in dimensions among the data and ensure comparability, this article adopts the mean-standard deviation method for standardization.

Step 2: Next, normalization is performed, followed by a correlation analysis of the indicators. Let r_{ij} represent the correlation coefficient between the original variables x_i and x_j. The correlation can be calculated using the formula:

$$R = \frac{Z^T Z}{n-1} = \left(r_{ij}\right)_{p\times p},$$

Among $r_{ij} = \frac{1}{n-1}\sum_{t=1}^{n} Z_{ti} Z_{tj} \left(i, j = 1, 2, ..., p\right)$, This step helps us understand the degree of correlation among various indicators.

Step 3: Based on the correlation analysis, we solve for the eigenvalues and eigenvectors of the correlation coefficient matrix to obtain the eigenvalues $\lambda_1, \ldots, \lambda_p$ of the correlation coefficient matrix R , and the corresponding eigenvectors are $\alpha_1, \alpha_2 \ldots, \alpha_p$, $\| \alpha_i \| = 1$, reach $\sum_{j=1}^{p} \alpha_{ij}^2 = 1$ among where α_{ij} is the j TH value of the vector α_j .

Step 4: Next, we proceed to calculate the variance contribution rate ϕ_i. The variance contribution rate serves as a valuable metric that aids us in comprehending the role played by each principal component in explaining the data information. The formula for its computation is as follows:

$$\phi_i = \frac{\lambda_i}{\sum_{k=1}^{p} \lambda_k},$$

The greater the variance contribution rate, the stronger the information capacity of the principal component synthesis variable.

Step 5: After obtaining the variance contribution rates, we then proceed to determine the factor contribution rates by calculating β_i . The formula for this computation is as follows:

$$\beta_i = \sum_{k=1}^{i} \frac{\lambda_k}{\sum_{k=1}^{p} \lambda_k},$$

Furthermore, we can determine which principal components are able to explain the majority of the data information. When β_i reaches over 80%, we can consider it as a principal component and accordingly determine the weights of the factors.

Step 6: Finally, we calculate the loadings of the principal components and perform a principal factor analysis. The relationship between the feature vector α_i and the principal component loadings l_{ij} is expressed as follows:

$$l_{ij} = \frac{\alpha_{ij}}{\sqrt{\lambda_i}},$$

This step helps us to gain a deeper understanding of the actual meanings represented by the principal components, as well as the degree of contribution of each variable in forming these principal components [7].

2.3. Quantification of Indicators for the Factors Influencing Residents' Happiness in Smart Cities

In recent years, as the pace of smart city construction has accelerated, research on the evaluation of smart cities has also gradually deepened. As mentioned in Section 1.2, there is an increasing number of evaluation studies in this field, and the research on happiness evaluation is also increasing day by day. The evaluation indicators and principles also show a diversified trend. Based on thorough preparation and a deep understanding of the development of smart cities, this article extracts the following core elements: happiness in smart ecological environment, happiness in smart economy, happiness in smart public services, happiness in smart politics, and happiness in smart safety management[8]. These five dimensions constitute the evaluation index system of the influencing factors of residents' happiness in smart cities. The specific meanings are shown in Table 2.

In the process of quantifying variables in this paper, a five-point scoring method was adopted. When measuring different degrees of values, the questionnaire set five answer options, where a higher score indicates a stronger willingness. Specifically, if a respondent chooses "Completely Agree" to a certain question, they will be assigned 5 points; if they choose "Somewhat Agree," they will be assigned 4 points; if they choose "Neutral," they will be assigned 3 points; if they choose "Somewhat Disagree," they will be assigned 2 points; and if they choose "Completely Disagree," they will be assigned 1 point. Through this quantification method, this paper evaluated questions related to the five dimensions in order to more accurately understand and analyze the data.

Table 2. Evaluation system

Evaluation Indicators	Specific Meanings
Wisdom Economy Happiness	Variables that measure residents' happiness from economic perspectives such as income and prices
Wisdom Politics Happiness	Variables that measure residents' happiness from political perspectives such as the government and its policies
Wisdom Public Service Happiness	Variables that measure the impact of public services such as healthcare and employment security on residents' happiness
Wisdom Safety Management Happiness	Variables that measure the impact of safety management services such as the internet, food, and transportation on residents' happiness
Wisdom Ecological Environment Happiness	Variables that measure the impact of ecological environments such as environmental protection and energy conservation on residents' happiness.

2.4. Comprehensive Evaluation of Factors Influencing Residents' Happiness in Smart Cities Based on Principal Component Analysis

2.4.1. Evaluation Indicators of Influencing Factors

Table 3. Indicators of influencing factors

Evaluation objective	Primary index	Secondary index
Influencing factors of residents' happiness under smart city	Intelligent government affairs	The implementation of relevant guiding policies by the Government [9](1)
		The extent to which the Government considers the needs of the public in the provision of smart services [9](2)
		The degree to which policy implementation procedures are fair, transparent and lawful (3)
	Intelligent public service	The ease of operation of intelligent service devices (4)
		The degree of rich and complete intelligent service content (5)
		The professionalism of the relevant service personnel (6)
		The responsibility of the relevant service personnel (7)
		How many intelligent service channels (8)
		The formulation and implementation of relevant systems(9)
	Intelligent security management	The extent to which residents are assisted in a timely manner when they encounter difficulties in using them(10)
		Whether the residents' right to intelligent services is protected by law[9](11)
		Availability of intelligent service information (12)
	Intelligent ecological environment	Reasonable situation of urban greening planning (13)
		Traffic congestion (14)
	Smart economy	Smart city talent construction situation (15)
		Whether the provision of intelligent services has stable financial support[10](16)
		Support degree of Internet of Things, cloud computing and other technologies[11](17)

Note: The (i) in the table refers to the component number for subsequent content, where $i = 1, 2, 3, \ldots 17$

Based on the research summarized in Section 1.2, this paper identifies the sources of residents' happiness in smart cities, including smart safety management, public services, ecological environment, government affairs, and economy. Through further exploration and research, the above five indicators are designated as primary indicators. Subsequently, 17 secondary indicators are established, such as the government's implementation of relevant guiding policies [9], the availability of smart service information, the number of service channels, traffic congestion, and the level of support provided by technologies such as the Internet of Things and cloud computing. These indicators are summarized in Table 3

2.4.2. KMO and Bartlett's Test

KMO test is an assessment tool that measures whether the index sample data is suitable for principal component factor analysis. Only when the KMO value exceeds 0.6, it indicates that the data is suitable for using the principal component factor analysis method. In testing the sample data of this article, we found that the KMO value is 0.966, thus proving that the sampling data in this article is suitable for conducting principal component factor analysis. Additionally, the result of Bartlett's test of sphericity also indicates that the significance value of the sample data is less than 0.05, further verifying the reliability of the sample data in this article from a statistical perspective, as shown in Table 4.

Table 4. KMO test and Bartlett test

KMO test and Bartlett test		
KMO value		0.966
Bartlett sphericity test	Approximate chi-square	1868.662
	df	136
	p	0.000***

Note: *** represents a significance level of 1%.

The result of the KMO test showed a KMO value of 0.966. Meanwhile, the outcome of the Bartlett's test of sphericity indicated a significant P-value (0.000***), suggesting that there is a significant difference between the observed results and the null hypothesis. Therefore, the null hypothesis is rejected, indicating that there is a correlation among the variables, and the principal component analysis is valid.

2.4.3. Results of Principal Component Analysis

Table 5. Principal component variance interpretation table

Element	Rotational front difference interpretation rate			Sum of squares of load after rotation		
	Aggregate	Percent variance of initial eigenvalue	Accumulative (%)	Aggregate	Variance percentage	Accumulative (%)
1	9.593	56.431	56.431	2.444	14.379	14.379
2	0.779	4.582	61.013	2.277	13.395	27.774
3	0.706	4.153	65.166	1.980	11.647	39.421
4	0.637	3.750	68.916	1.929	11.345	50.765
5	0.594	3.493	72.409	1.683	9.901	60.666
6	0.554	3.257	75.666	1.572	9.246	69.911
7	0.518	3.045	78.712	1.496	8.80	78.712
8	0.479	2.817	81.529			
9	0.459	2.70	84.229			
10	0.434	2.551	86.780			
11	0.398	2.341	89.122			
12	0.373	2.192	91.314			
13	0.345	2.027	95.241			
14	0.323	1.90	97.009			
15	0.301	1.768	98.631			
16	0.276	1.622	100.000			
17	0.233	1.369				

Note: The eigenvector matrix can reflect the load of each index on each principal component

Based on the aforementioned 17 quantitative influencing factors and evaluation indicators, a principal component analysis was conducted by setting the eigenvalue greater than 0.5 to obtain the contribution rate of each principal component, the rotated component matrix, and the eigenvector matrix. As shown in Table 2.4, the principal component analysis selected 7 principal components with contribution rates of 56.431% for the first principal component, 4.582% for the second, 4.153% for the third, 3.75% for the fourth, 3.493% for the fifth, 3.257% for the sixth, and 3.045% for the seventh. The cumulative contribution rate is 78.712%, indicating that the seven principal components cover over 78.712% of the information data from the 17 evaluation indicators, as shown in Table 5.

The first principal component is mainly determined by the government's implementation of relevant guiding policies in smart government administration. The second principal component pertains to the formulation and execution of relevant systems in smart government administration. The third principal component represents the rationality of urban greening, which belongs to the smart ecological environment. The fourth principal component is the level of support provided by technologies such as the Internet of Things and cloud computing, which falls into the category of smart economic services. The fifth principal component is composed of the professional quality of relevant service personnel in smart public services. The sixth principal component primarily concerns traffic congestion, which belongs to the aspect of the smart ecological environment. The seventh principal component represents the number of service channels, which falls into the category of smart public services. As shown in Table 6.

Table 6. Component matrix after rotation

Designation	Element						
	1	2	3	4	5	6	7
The implementation of relevant guiding policies by the Government [9]	0.744	0.091	0.191	0.317	0.255	0.227	0.163
Availability of information about smart services	0.386	0.219	0.416	0.358	0.48	0.094	0.044
The number of service channels	0.199	0.191	0.133	0.281	0.091	0.152	0.826
The ease of operation of intelligent service devices	0.227	0.672	0.278	0.327	0.227	0.083	0.15
The degree of rich and complete intelligent service content	0.456	0.404	0.28	0.109	0.386	0.274	0.15
The professionalism of the relevant service personnel	0.162	0.263	0.133	0.419	0.701	0.277	0.034
The responsibility of the relevant service personnel	0.610	0.524	0.131	0.274	0.146	0.055	0.184
The extent to which the Government prioritizes the interests and needs of the public in the provision of smart services [9]	0.610	0.321	0.254	0.117	0.145	0.296	0.252
The extent to which residents are assisted in a timely manner when they encounter difficulties in using them	0.310	0.196	0.171	-0.058	0.633	0.135	0.522
Whether the provision of intelligent services has stable financial support [10]	0.152	0.463	0.62	0.094	0.163	0.118	0.301
Whether the right of residents to enjoy smart services is protected by law [9]	0.391	0.361	0.276	0.245	0.154	0.42	0.265
Support degree of Internet of Things, cloud computing and other technologies [11]	0.244	0.263	0.161	0.721	0.132	0.196	0.165
Smart city talent construction situation	0.215	0.144	0.456	0.581	0.182	0.109	0.32
The degree to which policy implementation procedures are fair, transparent and lawful	0.388	0.406	0.206	0.436	0.198	0.275	0.164
The formulation and implementation of relevant systems	0.214	0.663	0.135	0.172	0.222	0.442	0.175
Traffic congestion	0.23	0.17	0.223	0.187	0.194	0.808	0.133
Reasonable situation of urban greening planning	0.211	0.124	0.814	0.23	0.128	0.238	0.05

In this paper, the standardized data of 17 variables, including traffic congestion, the extent to which the government considers public needs in providing smart services, the formulation and implementation of relevant policies, the level of support provided by technologies such as the Internet of Things and cloud computing[11], the rationality of urban greening planning, the government's implementation of relevant guiding policies[9], the development of smart city talent, the number of smart service channels, the reasonableness of policy implementation procedures, the sense of responsibility of relevant service personnel, the comprehensiveness of smart service content, the professional quality of relevant service personnel, the extent to which residents receive timely assistance when encountering difficulties, the availability of smart service information, the stability of funding support for the provision of smart services[10], the legal protection of residents' rights to smart services[9], and the ease of operation of smart service equipment, are sequentially labeled as $x_1 \sim x_{17}$. Assuming that the first principal component to the seventh principal component are sequentially labeled as F_1, F_2, F_3, F_4, F_5, F_6, F_7, then according to Table 7, the principal component expressions are obtained as follows:

$$\begin{aligned}F_1 = &0.744x_1 + 0.386x_2 + 0.199x_3 + 0.227x_4 + 0.456x_5 + 0.162x_6 \\ &+0.610x_7 + 0.610x_8 + 0.310x_9 + 0.152x_{10} + 0.391x_{11} + 0.244x_{12} \\ &+0.215x_{13} + 0.388x_{14} + 0.214x_{15} + 0.230x_{16} + 0.211x_{17}\end{aligned},$$

$$\begin{aligned}F_2 = &0.091x_1 + 0.219x_2 + 0.191x_3 + 0.672x_4 + 0.404x_5 + 0.263x_6 \\ &+0.524x_7 + 0.321x_8 + 0.196x_9 + 0.463x_{10} + 0.361x_{11} + 0.263x_{12} \\ &+0.144x_{13} + 0.406x_{14} + 0.663x_{15} + 0.17x_{16} + 0.124x_{17}\end{aligned},$$

$$\begin{aligned}F_3 = &0.191x_1 + 0.416x_2 + 0.133x_3 + 0.278x_4 + 0.28x_5 + 0.133x_6 \\ &+0.131x_7 + 0.254x_8 + 0.171x_9 + 0.62x_{10} + 0.276x_{11} + 0.161x_{12} \\ &+0.456x_{13} + 0.206x_{14} + 0.135x_{15} + 0.223x_{16} + 0.814x_{17}\end{aligned},$$

$$\begin{aligned}F_4 = &0.317x_1 + 0.358x_2 + 0.281x_3 + 0.327x_4 + 0.109x_5 + 0.419x_6 \\ &+0.274x_7 + 0.117x_8 - 0.058x_9 + 0.094x_{10} + 0.245x_{11} + 0.721x_{12} \\ &+0.581x_{13} + 0.436x_{14} + 0.172x_{15} + 0.187x_{16} + 0.230x_{17}\end{aligned},$$

$$\begin{aligned}F_5 = &0.255x_1 + 0.48x_2 + 0.091x_3 + 0.227x_4 + 0.386x_5 + 0.701x_6 \\ &+0.146x_7 + 0.145x_8 + 0.633x_9 + 0.163x_{10} + 0.154x_{11} + 0.132x_{12} \\ &+0.182x_{13} + 0.198x_{14} + 0.222x_{15} + 0.194x_{16} + 0.128x_{17}\end{aligned},$$

$$\begin{aligned}F_6 = &0.227x_1 + 0.094x_2 + 0.152x_3 + 0.083x_4 + 0.274x_5 + 0.277x_6 \\ &+0.055x_7 + 0.296x_8 + 0.135x_9 + 0.118x_{10} + 0.42x_{11} + 0.196x_{12} \\ &+0.109x_{13} + 0.275x_{14} + 0.442x_{15} + 0.808x_{16} + 0.238x_{17}\end{aligned},$$

$$\begin{aligned}F_7 = &0.163x_1 + 0.044x_2 + 0.826x_3 + 0.15x_4 + 0.15x_5 + 0.034x_6 \\ &+0.184x_7 + 0.252x_8 + 0.522x_9 + 0.301x_{10} + 0.265x_{11} + 0.165x_{12} \\ &+0.32x_{13} + 0.164x_{14} + 0.175x_{15} + 0.133x_{16} + 0.05x_{17}\end{aligned}.$$

Then, after calculating the above 7 principal components with the quantified variable value, taking the variance contribution rate of each principal component as the weight, the comprehensive evaluation index of influencing factors of residents' happiness under smart city is constructed as follows:

$$\begin{aligned}Y = &0.56431 \times F_1 + 0.04582 \times F_2 + 0.04153 \times F_3 + 0.0375 \times F_4 \\ &+0.03493 \times F_5 + 0.03257 \times F_6 + 0.03045 \times F_7\end{aligned}$$

Through the principal component analysis, we can conclude that the main factors influencing residents' happiness in smart cities are the government's implementation of relevant guiding policies, traffic congestion, the extent to which the government considers public needs in providing smart services, and the rationality of urban greening planning. Among these factors, the government's implementation of relevant guiding policies accounts for the highest proportion, indicating that it is currently the most significant aspect affecting residents' happiness in smart cities. Therefore, the government should prioritize addressing this issue.

3. An Empirical Analysis of the Influencing Factors of Residents' Happiness in Smart Cities by Random Forest Model

3.1. Symbol Description

The symbols in the random forest model are described below, as shown in Table 7.

Table 7. Description of symbols

Symbol	Explain
Θ_k	Parameter set
$h_i(x)$	Decision tree classification model
$h(X, \Theta_k), k = 1, \ldots$	Classification tree model
$h_1(x), h_2(x), \ldots, h_k(x)$	Classification model
$H(x)$	Combinatorial classification model
Y	Objective function

3.2. Classification Principle of Random Forest Model

Random forest (RF) is a powerful statistical learning technique whose core idea is to build multiple decision trees by drawing different subsets of samples from the original data set multiple times. These decision trees do not work independently, but through a voting mechanism, the predictions of each tree are combined to arrive at a final prediction. This method not only significantly improves the accuracy of the prediction, but also has excellent tolerance for outliers and noise in the data. More importantly, random forest can effectively avoid the occurrence of overfitting phenomenon, that is, it can maintain good generalization performance even when the model complexity is too high.

Random forest classification (RFC) is an ensemble learning technique that improves classification performance by combining the predictions of multiple decision tree classification models $\{h(X, \Theta_k), k = 1, \ldots\}$. By integrating the outputs of multiple models (i.e., decision trees), this combined approach reduces the risk of overfitting that can arise from a single model and increases overall generalization. Its core lies in the fact that the construction of each decision tree model is based on an independent and equally distributed set of random vector parameters $\{\Theta_k\}$, which enables each model to provide unique and valuable information in prediction [12].

When given independent variables, each decision tree has equal voting rights, and they jointly decide the final classification result through the majority voting mechanism. The core idea of random forest classification (RFC) is to firstly use Bootstrap sampling technology to randomly extract multiple sub-samples with the same size as the original data set from the original training set. Then, for each subsample, an independent decision tree model is built to get a variety of different classification results. Finally, by aggregating these independent classification results, a voting mechanism is used to determine the final category belonging of each sample.

This method not only significantly improves the accuracy and robustness of classification, but also shows the powerful processing ability of noise and outliers, which makes the random forest classification method more advantageous in practical application. At the same time, due to the idea of ensemble learning, RFC can also avoid the problem of overfitting to a certain extent, so that the model has better generalization ability in the face of new data, as shown in Figure 1.

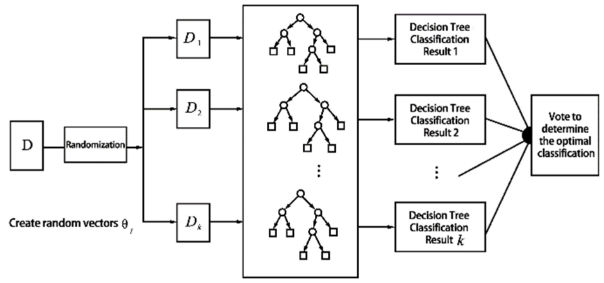

Figure 1. Schematic diagram of a random forest

Random Forest (RF) enhances the diversity among individual classification models by generating multiple different training sets, thus improving the predictive ability of the combined classification model on unknown data. This process typically involves repeating the training process k times, resulting in the generation of an independent decision tree classification model $h_i(x)$ in each round, which forms a sequence of classification models $\{h_1(x), h_2(x), ..., h_k(x)\}$. Due to the use of randomization methods (such as feature selection and sample selection) during the construction of these models, there will be a certain degree of variation among them [12], which is precisely the core of Random Forest.

After obtaining this series of classification models, they collectively constitute an efficient multi-classification model system. When faced with new data, the system inputs each data sample into all the decision tree models, and each model independently provides its classification insights. Finally, by applying a simple majority voting method, the system selects the result that is most commonly identified by the models as a particular category as the final classification decision, ensuring that each classification is more accurate and reliable. The calculation method is as follows:

$$H(x) = \arg\max_{Y} \sum_{i=1}^{k} \mathrm{I}\left(h_i(x) = Y\right),$$

In this context, $H(x)$ represents the ensemble classification model, h_i is a single decision tree classification model, Y is the output variable (or target variable), and $I(x)$ is an indicator function.

3.3. Ranking the Importance of Factors Using Random Forest Classification

Table 8. Ranking table of the importance of random forest for influencing factors of residents' happiness

Influence factor	Degree of importance
The implementation of relevant guiding policies by the Government [9]	0.135653
The extent to which the Government considers the needs of the public in the provision of smart services[9]	0.118453
The degree to which policy implementation procedures are fair, transparent and lawful	0.052313
The ease of operation of intelligent service devices	0.012296
The degree of rich and complete intelligent service content	0.023544
The professionalism of the relevant service personnel	0.043175
The responsibility of the relevant service personnel	0.090775
How many intelligent service channels	0.059852
The formulation and implementation of relevant systems	0.046568
The extent to which residents are assisted in a timely manner when they encounter difficulties in using them	0.029325
Whether the residents' right to intelligent services is protected by law [9]	0.02392
Availability of intelligent service information	0.036928
Reasonable situation of urban greening planning	0.078531
Traffic congestion	0.06786
Smart city talent construction situation	0.058741
Whether the provision of intelligent services has stable financial support [10]	0.036587
Support degree of Internet of Things, cloud computing and other technologies [11]	0.085479

This article aims to establish a random forest classification model to explore the impact of 17 influencing factors (independent variables X_i) on residents' sense of happiness (dependent variable Y) established in section 2.1, including the situation of government-issued relevant guiding policies for fulfillment and development, the extent to which the government considers public needs in providing intelligent services, whether residents' rights to intelligent services are legally protected, the degree of fairness, transparency, legality, and rationality of policy implementation procedures, the ease of operation of intelligent service equipment, the stability of financial support for the provision of intelligent services, the completeness and comprehensiveness of intelligent service content, the professional quality of relevant service personnel, the sense of responsibility of relevant service personnel, the number of intelligent service channels, the formulation and implementation of relevant systems, the availability of intelligent service information, the rationality of urban greening planning, traffic congestion, smart city talent development, the extent to which residents can promptly receive help when encountering difficulties in usage, and the support degree of technologies such as the Internet of Things and cloud computing. The goal is to construct a regression tree model.

To present the data in a more intuitive and vivid manner, this article will transform Table 8 into a bar chart. This transformation will not only make the distribution and comparison of data readily apparent, but also highlight the differences among various data, further enhancing our understanding of the data's implications and trends. Through the bar chart display, we can more clearly see the degree of importance of each data point, as shown in Figure 2.

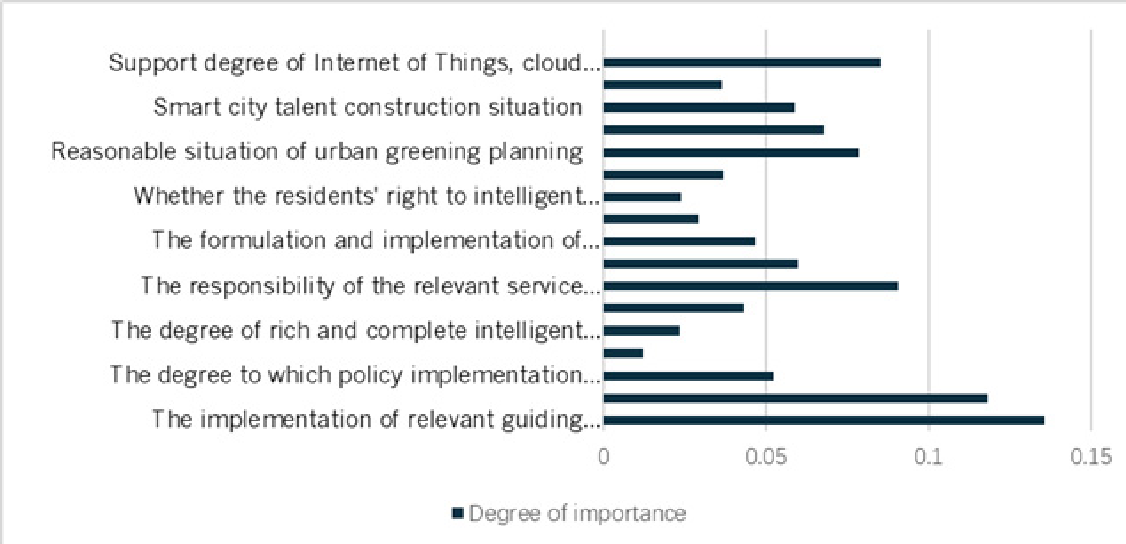

Figure 2. Ranking diagram of random forest importance of factors affecting residents' happiness

Based on Table 8 and Figure 2, we can see that the government's implementation of relevant guiding policies, the extent to which the government considers public needs in providing intelligent services, the formulation and implementation of relevant systems, the sense of responsibility of relevant service personnel, and the support degree of technologies such as the Internet of Things and cloud computing have a significant impact on residents' sense of happiness. It can be observed that the results obtained from the random forest model are roughly consistent with those obtained from the principal component analysis, indicating the credibility of the analysis results.

In conclusion, the government should actively take measures to further enhance the rationality of policy implementation. In providing intelligent services, the government should deeply consider public needs to ensure that the services are more in line with people's actual needs. At the same time, the government and enterprises should increase research and development efforts on technologies such as the Internet of Things and cloud computing, continuously promoting technological innovation and application. In addition, governments at all levels should strengthen the rationality review of urban greening planning, taking into account local conditions and avoiding the situation of "a mandarin orange grown in the south tastes different when transplanted to the north." They should ensure that greening construction meets local ecological environments and people's needs. Through these measures, the government will be able to better promote social progress and development, enhancing residents' sense of happiness.

Acknowledgments

2022 National College Student Innovation Training Program Project "Research on Evaluation and Improvement Measures of Residents' Happiness under Smart City" (202210378208).

References

[1] Wang Peng, Jia Yinghui, Li Qiushuang. Research on the development strategy of Smart Beijing from the perspective of digital economy benchmark city construction [J]. Digital Economy, 2022, 9(08): 8-14.

[2] Vijendcr K. S., Somcsh K., Vijay B. S. Advanced automated module for smart and secure city[J]. Procedia Computer Science, 2016, 7(78): 367-374.

[3] Cottrel, S. P, Vaske J. J., Roemer, J. M. Resident satisfaction with sustainable tourism: The case of Franken Wald Nature Park, Germany[J]. Tourism Management Perspectives, 2013, 2(8): 42-48.

[4] Bjonskov, Christian, Axel Drehe. Formal institutions and subjective Wel-being: Revisiting the cross-country evidence[J]. European Journal of Political Economy, 2010, 119(3): 001.

[5] Mccrea R, Stlmson R, Western J. Testing a moderated model of satisfaction with urban living using data for Brisbane-South East Queensland, Australia[J]. Social Indicators Research, 2005, 32(2): 121-152.

[6] Zhou Ganghui. Research on branding and communication strategies of Smart City [D]. Hangzhou Normal University. 2022.

[7] Qiu Yixin. Application of Principal Component Analysis-Fuzzy Comprehensive Evaluation Model in Regional Ambient

Air Quality Evaluation [J]. Modern Chemical Research, 2023, (24): 130-132.

[8] Wan Shaojun. Research on demand-oriented thematic evaluation index system for Smart City [D]. Nanchang University, 2023.

[9] Wang Yan. Research on Residents' Sense of Acquisition Evaluation and Improvement Path of Smart Cities [D]. Southeast University, 2020.

[10] Ren Shuying. Research on the development of Smart cities [D]. Shenyang Normal University, 2015.

[11] Miao Yanxu. Exploring the application of Internet of Things and cloud computing technology in smart city [J]. Technological innovation and Application, 2020, 10(18): 165-166.

[12] Fang Quannan, WU Jianbin, ZHU Jianping. A review of stochastic forest methods [J]. Statistics and Information Forum, 2011, 26(03): 32-38.