ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

Research on the Impact of Hangzhou's Allocation Policies on School District Housing Based on Machine Learning Algorithms

Yihan Ye[1], Xin Zhou[1,a,*], Peiying Zheng[1], Yujie Han[1], Xinsheng Ma[1]

[1]Zhejiang International Studies University, Hangzhou, China
[a]echozhoux@gmail.com
**Corresponding author*

Abstract: *This paper focuses on the impact of Hangzhou's allocation policies on school district housing. Using a multistage sampling method and stratified sampling method, the researchers selected a sample size and survey subjects from the original six districts in Hangzhou. A structured questionnaire and Likert scale method were used to design the survey questionnaire, and in-depth interview questions were designed based on the principles of in-depth interviews. After completing the distribution and acquisition of the questionnaire and the cleaning of the relevant data, the researchers used the multiple logistic regression model to study the significance of the eight factors that affect parents' intention to buy a house, and then used various machine learning classification methods such as KNN, decision tree, Boosting, Bagging, and random forest, and calculated the influence of the influence variables on the intention to buy a house using 10-fold cross-validation. Based on this analysis, housing price differences in the three months before and after the policy were analyzed, and related policy recommendations were put forward.*

Keywords: *Machine Learning Algorithms, Allocation Policies, School District Housing*

1. Introduction

In April 2021, the Hangzhou Municipal Bureau of Education issued the "Notice on 2021 Hangzhou Municipal Enrollment Work for Various High Schools in the Urban Area" (hereinafter referred to as the Allocation Policy) ".The introduction of the new policy is bound to have an impact on parents' intention to buy a house, and this impact will affect the rise and fall of housing prices in Hangzhou School District. In this way, it will have a great impact on the "school district housing boom" in Hangzhou, and may even set off a huge wave in China's real estate sector, which will have a profound impact on China's education resource allocation. This study will used various machine learning classification methods to study the "cooling" effect of the new distribution policy in Hangzhou on the phenomenon of "school district housing fever", test the implementation effect of the new distribution policy in Hangzhou, and also promote the distribution policy to a certain extent, promote the equal distribution of resources, and provide a useful reference for urban planning.

2. Research Background

In recent years, with the rapid advancement of urbanization in Hangzhou, there has been a tremendous transformation in the urban social space[1]. To promote property appreciation, real estate companies introduce famous schools near new housing estates, which leads to a doubling of the prices of school district houses, hot sales, and old housing estates unable to meet demand, with prices reaching new highs and even the phenomenon of "high price, few houses," "priced, no houses," and "high school district houses." The high demand for school district houses reflects the imbalanced distribution of educational resources and the insufficient investment in education[2].

The "Notice on 2021 Hangzhou Municipal Enrollment Work for Various High Schools in the Urban Area" (hereinafter referred to as the Allocation Policy) issued in March 2021 stipulates that key provincial general high schools, provincial general high school characteristic demonstration schools, and their affiliated schools will enroll students with allocation quotas. As a reform attempt by the Hangzhou Municipal Government and Education Bureau to implement educational equality, this policy quickly became a hot topic in the education sector. At present, research on the "hot school district house"

ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

phenomenon generally focuses on the causes of the "hot school district house" phenomenon, the impact of the "hot school district house" phenomenon[3], and strategies for controlling the "hot school district house" phenomenon. However, the new Allocation Policy in Hangzhou has just been released and relevant research has not yet been carried out promptly. Whether the Allocation Policy will create waves in the "feverish" school district house field, whether it will be beneficial to the equalization of education and public services, and whether it can reduce the contradiction between education and residential space remains to be seen. This project will base on the work of previous research, abandon traditional information and data processing algorithms and adopt machine learning algorithms to conduct a comprehensive analysis of the Allocation Policy's impact on the "hot school district house" phenomenon, and provide policy recommendations for improving the allocation of educational resources and reducing the "hot school district house" phenomenon.

3. Theoretical Framework

This project plans to acquire data through questionnaire surveys, apply statistical modeling and machine learning algorithms, and study the factors that influence the intention of different groups of parents to purchase school district houses under the allocation policy, as well as the influence on the prices of school district houses in different areas.

For this purpose, the project takes parents and school district house prices within the scope of the new allocation policy as the survey subjects, takes the original six districts of Hangzhou as the survey scope, and selects several families in each region based on multi-level sampling. Using structured questionnaires and the Likert scale method, the questionnaire was designed based on the principles of in-depth interviews to design interview questions, and research on school district houses was conducted through on-site visits, online research, and other methods to obtain research data and establish relevant statistical survey models to support the theory. Based on this, the views and purchasing intentions of different parents on school district houses are summarized to determine whether the implementation of this new allocation policy can substantially improve the "high temperature" of school district houses and the uneven distribution of educational resources and spatial distribution in Hangzhou.

This study distributed a total of 300 questionnaires and received 237 valid questionnaires, for a recovery rate of 79.0%. The sample design was effective and consistent with reality. Using SPSS to calculate the reliability of the survey questionnaire, Cronbach's Alpha value was obtained as 0.750, and the standardized Cronbach's Alpha value was 0.746, with these two coefficients approaching 80%, indicating that the data collected by this survey questionnaire has a high degree of internal consistency and reliability. Using SPSS's factor analysis function, the KMO value was calculated as 0.763, the Bartlett value as 317.938, and the degrees of freedom as 55, all of which exceed the critical value, indicating that the data is suitable for factor analysis. The factor analysis extracted three factors: "parental education background," "family economic conditions," and "housing preferences." The variance contribution rate of the three factors is 40.9%, which shows that these three factors have a significant impact on the purchasing intention of school district houses. The regression analysis shows that the "parental education background," "family economic conditions," and "housing preferences" factors have a significant positive impact on the purchasing intention of school district houses.

4. Research Resluts

4.1. Multivariate Logistic Model Solution and Analysis

Using the multivariate logistic regression model, the parent's intention to buy a house (yes or no) is taken as the dependent variable and 8 influencing factors are taken as the independent variables. Using SPSS, the results are shown in Tables 1-3."[4]

From Table 1, it can be obtained that the global test value is <0.05, that is, the model is statistically significant; according to the three pseudo-determination coefficients, the model is within the acceptable range.

ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

Table 1: Model fitting information.

	-2 log-likelihood	Bangla	df	significant
intercept only	256.156			
finally	142.317	113.839	16	0.000
Cox and Snell		0.831		
Nagelkerke		0.513		
McFadden		0.353		
	-2 log-likelihood	Bangla	df	significant
intercept only	256.156			
finally	142.317	113.839	16	0.000
Cox and Snell		0.831		
Nagelkerke		0.513		
McFadden		0.353		

Table 2: Likelihood ratio test table.

effect	Model Fit Criteria	likelihood ratio test		
	-2-fold log-likelihood of the reduced model	Bangla	df	significant level
Parent education	143.286	.969	2	.616
Schools admitted after the high school entrance examination	143.282	.966	1	.326
Usual grades	146.944	4.627	2	.099
Learn degree	148.155	5.838	2	.054
quota acquisition	155.834	13.517	2	.001
educational equity	152.201	9.884	1	.002
average monthly income	197.190	54.874	4	.000
house price judgment	143.114	.797	2	.671

The likelihood ratio test is shown in Table 2. It can be seen that the parameter test of parents' education, parents' expectation of their children's whereabouts after the high school entrance examination, and children's usual grades are not significant, so they are not used as independent variables of the model. They are recalculated using SPSS. For the parameter estimation table and significance test, see table 3.[5]

Table 3: Parameter estimation table of influencing factors for parents to buy a house.

		coefficient	standard error	Wald	significant
Usual grades	not good	-2.169	1.000	4.700	.030
	medium	-.479	.503	.908	.341
	excellent	0b	.	.	.
Learn degree	do not know	1.686	.698	5.838	.016
	general know	.319	.507	.397	.529
	know very well	0b	.	.	.
quota acquisition	difficulty	1.819	.662	7.539	.006
	generally	-.158	.572	.076	.783
	easy	0b	.	.	.
educational equity	No	1.169	.391	8.950	.003
	Yes	0b	.	.	.
average monthly income	<4500	4.348	1.558	7.790	.005
	4500~9000	3.244	1.158	7.846	.005
	9000~35000	.881	1.130	.608	.436
	35000~55000	-.058	1.186	.002	.961
	>55000	0b	.	.	.

It can be seen from Table 3 that at the 0.05 significance level, those with poor grades, ignorance of

ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

student allocation policies, difficulties in obtaining student allocation quotas, thinking that student allocation policies are not conducive to educational equity, monthly income < 4500, and monthly income between 4500 and 9000 There is a significant correlation with parents' intention to buy a house. It can be seen that parents who do not understand the allocation student policy, think that their children are less likely to obtain an allocation student quota, think that the allocation student policy is not conducive to education fairness, and parents with low monthly income have low intentions to buy a house, while parents with high monthly income have a high intention to buy a house.

4.2. Calculation of Machine Learning Classification Algorithm and Analysis of Cross-Validation Results

Through R software programming and calculation, the classification accuracy of five machine learning classification algorithms and 10-fold cross-validation is shown in Table 4.[6]

Table 4: Classification accuracy of five machine learning classification algorithms

classification algorithm	classification accuracy	10-fold cross-validation classification accuracy
KNN algorithm	0.82	0.651
Boosting algorithm	0.86	0.735
Decision tree algorithm	0.81	0.708
Bagging algorithm	0.81	0.729
random forest algorithm	0.83	0.752

According to Table 4, when all the data are used for prediction, the five classification algorithms all have good classification accuracy, and the Boosting algorithm, Random Forest algorithm and KNN algorithm have higher accuracy, which are 0.86, 0.83 and 0.82 respectively. Using 10-fold cross-validation to recalculate, the classification accuracies of the three methods are 0.735, 0.651 and 0.752, respectively. It can be seen that classification accuracy of the Random Forest algorithm is the highest. Therefore, the Random Forest algorithm is used to analyze the importance of the eight influencing factors, and the results are shown in Figure 1.[7]

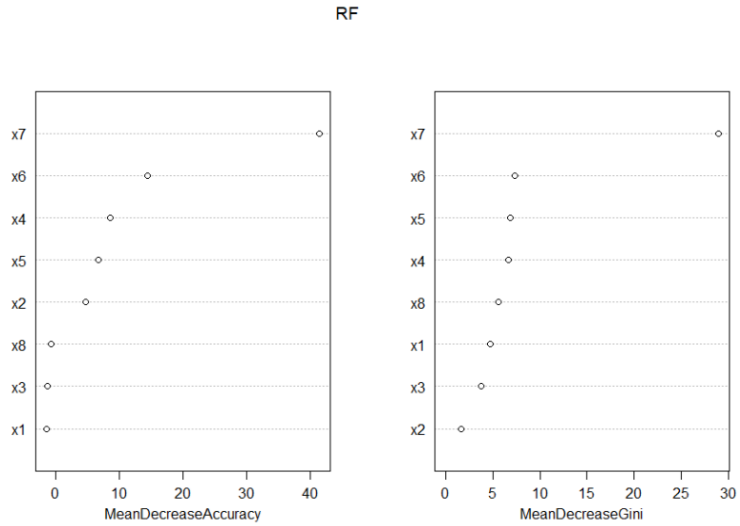

Figure 1: Mean Decrease Accuracy and Mean Decrease Gini coefficient graph.

From the Mean Decrease Accuracy and Mean Decrease Gini coefficients in Figure 1, it can be seen that the importance from high to low, the two most important variables are average monthly income and educational equity, followed by policy understanding and the possibility of obtaining quotas. The destination of the senior high school entrance examination, house price judgment, usual grades, and parents' education are the four variables with the lowest degree of influence.[8]

5. Research Proposals

5.1. Increase Publicity and Fully Understand the Policy

The survey results show that the new student allocation policy can effectively curb the phenomenon of "school district housing fever", but the policy publicity is insufficient. Parents' understanding of the

ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

new policy for the high school entrance examination is not high enough, which will allocate students' policy unable to exert its due effect. Improve the uneven distribution of educational resources to a greater extent. The government and education departments at all levels should strengthen the publicity of the New Deal, and use the media and online platforms to introduce the specific content of the New Deal. Schools can increase students' and parents' understanding of the New Policy through home-school communication, class meetings, and lectures on further education. They can also set up relevant question-and-answer channels in the school to provide personalized services for different students, answer students' doubts about further education, and ensure that middle school students and Parents make choices based on a comprehensive understanding of the New Policy on the allocation of students for senior high school entrance examinations, to provide each student with a better guarantee of education fairness.[9]

5.2. Comply with Current Political Trends, Rationally Analyze School Selection

Compared with the previous system of "recommended students" for the senior high school entrance examination, the policy of allocating students has changed the criteria for students to choose schools. High schools directly allocate quotas to each middle school according to the number of students in the school. This means that some ordinary middle school students with a large number of When the policy of "recommended students" is implemented, there is a greater chance of entering key high schools, thereby reducing the situation where key junior high school students monopolize key high school places. The release of the new student allocation policy represents the emergence of a new opportunity, but it does not mean that the policy is suitable for all candidates. Whether to participate in the competition for the number of allocated students requires careful study and careful consideration by parents and students, and fully weighs the probability and quality of the allocated students. Educational resources, housing prices in school districts and many other factors are used to choose a way that is more suitable for candidates to enter higher education.

5.3. Choose a School District Reasonably, Choose a House Rationally and Buy a House

Before the New Deal, the factors affecting housing prices in school districts were relatively concentrated. Under the same circumstances of other factors, the richer the educational resources in the school district, the higher the housing price in the school district. After the introduction of the student allocation policy, the factors affecting housing prices in school districts have become more complicated. The changes in the choice of educational resources by students and parents are likely to affect the housing market. Although most housing prices will not change significantly in the short term, for some children For parents who are still in the primary education stage, they should choose rationally and weigh the pros and cons when buying a school district house in the future. The level of school you choose will have a certain impact on whether your child will be able to obtain a student quota in the future. Therefore, when choosing to buy a school district house, you should take a long-term view. For those who buy houses not for living but only for investment, they should not blindly speculate in real estate or follow the trend of investment without understanding the market wind deflection brought about by policy changes. They may face great risks. Therefore, it is reasonable to choose school districts and be rational. Choosing a house is very important.

6. Conclusions

The research shows that the introduction of the distribution student policy has led to a certain downward trend in the real estate price of Hangzhou School District, but has not yet set off a huge wave in the real estate field. However, it is undeniable that the distribution student policy vigorously implemented by Hangzhou is a new attempt and initiative, and whether its publicity and influence reach the desired scale needs further research.

References

[1] Mao F F, Luo G F, Pan J S. Research on the Impact of High-Quality Educational Resources on Housing Prices in Hangzhou School District [J]. Urban and Environmental Research, 2014, 1(02): 53-64.

[2] Zhang Y. The Influence of Differentiation of Educational Resources on Real Estate Prices Taking Hangzhou West Lake District as an Example [D]. Hangzhou: Zhejiang University, 2015.

[3] Huang Y. Research on Housing Heat in School Districts [D]. Changsha: Hunan Normal University,

ISSN 2663-8169 Vol. 5, Issue 1: 80-85, DOI: 10.25236/IJNDE.2023.050114

2015.
[4] Liu J C, Zhang L L. Analysis on the Phenomenon of "Housing Fever in School Districts" [J]. Educational Science Forum, 2011(12): 61-63.
[5] Si S K, Sun X J. Mathematical Modeling Algorithms and Applications (3rd Edition) [M]. Beijing: National Defense Industry Press, 2021.
[6] He X Q. Multivariate Statistical Analysis (Fifth Edition) [M]. Beijing: Renmin University of China Press, 2019.
[7] Wang B H. Multivariate Statistical Analysis and R Language Modeling (Fifth Edition) [M]. Beijing: Higher Education Press, 2020.
[8] Wu X Z, Zhang M. Applied Regression and Classification—Realization Based on R and Python [M]. Beijing: Renmin University of China Press, 2020.
[9] Wei Z S. Probability Theory and Mathematical Statistics Course (Second Edition) [M]. Beijing: Higher Education Press, 2008.