Unraveling the relationship between the digital platforms and gentrification: Evidence from Nanjing, China

Qianhui He
Affiliation: Southeast University
Email: 230238006@seu.edu.cn

Shijie Sun
Affiliation: Southeast University
Email: sun_shijie@seu.edu.cn

Abstract
Information-driven digital platforms are reshaping urban and social space, often intensifying gentrification. Using Nanjing (2019-2024) as a case, we combine Beike second-hand housing transactions with Weibo check-ins to map community gentrification stages, applying spatial autocorrelation, cluster and regression analyses. Gentrification correlates strongly with check-in density: commercial cores, tourist spots and university districts form "high-high" clusters, while peripheral zones remain "low-low". Check-ins by non-locals, KOLs and KOCs exert greater influence than those by residents. Algorithmic "discursive investing" turns informational capital into place branding, attracting investors and accelerating change. These findings highlight digital platforms as new engines—and amplifiers—of urban inequality, calling for governance that regulates platform algorithms and protects spatial justice.

Keyword: gentrification-digital platforms-Social Media Check-ins; SHAP

Introduction

Gentrification was initially used to describe the socio-spatial transformation in inner-city London, where the middle class moved into declining working-class neighborhoods, revitalised the physical environment, and established residential enclaves(Glass, 1964). Over the past half-century, research on gentrification has expanded from its early focus on class displacement in urban neighborhoods of Western Europe and North America to encompass narratives of socio-spatial change across the Global South and broader urban-rural contexts. The "planetary gentrification" (Shen et al., 2018) theory posits that gentrification has become an inevitable experienced phase in global urban development, emphasizing the need to examine its processes across diverse geographical, economic, political, and cultural settings worldwide. In the context of the informatised city, many scholars have begun exploring the relationship between geo-media(Hartmann and Jansson, 2024), urban mediatisation(Törnberg and Uitermark, 2022), digital platforms(Shan et al., 2023), and planetary gentrification. Against the backdrop of accelerating global digital transformation, platform capitalism is reshaping the logic of urban spatial production through location-based services (LBS). Traditional gentrification studies emphasize the linear relationship between physical space renovation and class displacement but overlook the hidden power of digital platforms in shaping spatial perception through information flows.

China, as the world's largest social media application market, sees platforms like Weibo and Douyin generating over 500 million daily location tags. Social-consumption platforms such as Weibo, Dianping, and Xiaohongshu are reshaping the visibility and appeal of urban spaces through algorithmic recommendations and user-generated content. These platforms not only alter consumption patterns but also invisibly accelerate the demographic and landscape reconfiguration of specific neighborhoods, creating a new transmission chain of "digital footprints-spatial value-capital flow." Against this backdrop, this study takes Nanjing as a typical case to reveal the relationship between platform activities, user behavior, and gentrification, offering a new perspective for urban research in the Global South context.

Digital Platforms and Gentrification

Currently, Western scholars have begun exploring the relationship between gentrification and digital platforms, including social media platforms, lifestyle service platforms, and the sharing economy. Concepts such as platform capitalism and mediated narratives of gentrification have emerged as key theoretical tools for understanding gentrification in the new era. On one hand, cities and capital have always been deeply intertwined. Digital platforms have shifted capital's exploitation of urban space from explicit domination to implicit discipline. Through algorithmic "black box" operations and discursive construction, urban spaces are commodified as "idle" resources to facilitate capital accumulation. On the other hand, digital platforms reshape the terrain of identity formation. The middle class utilises these platforms as stages to showcase their aesthetic preferences, attracting like-minded groups and fostering spatial clustering. This leads to a more subtle form of cultural isolation and social segregation, where exclusion occurs under the guise of digital connectivity.

2.1 Digital Platform "Discourse" and Spatial Representation

Digital platforms possess the capacity to frame urban spaces as consumable commodities or investment opportunities, constructing specific discourses that shape spatial representation and attract capital to facilitate spatial valorisation. Törnberg et al. examined the portrayal of Vidigal across diverse media platforms-such as YouTube, Airbnb, and Flickr-demonstrating how these platforms strategically reimagined the neighborhood from a stigmatised, impoverished, and hazardous area into a dynamic space brimming with opportunity. By accentuating its "heterogeneity" and "authenticity," these platforms effectively drew in tourists, new residents, and investors. Consequently, new residential and commercial developments gradually eroded local cultural heritage and transformed community structures, accelerating the area's gentrification(Törnberg and Uitermark, 2022). Similarly, Erika Polson's case study of Denver's RiNo Art District analyzed the interplay between graffiti, Instagram, and gentrification. The study argues that social media platforms are reconfiguring the production, aesthetics, and consumption of graffiti and street art. The viral dissemination of Instagram imagery has repurposed street art as a tool for enhancing neighborhood cultural capital, making its commercial potential visible to upper-middle-class residents and investors. This process ultimately attracts capital inflows, further accelerating gentrification (Polson, 2024).

In a critical examination of Instagram and Facebook posts related to Brisbane's West End, Walters highlights the hegemonic influence of social media in constructing immaterial yet highly symbolic spatial narratives. The study contends that digitally mediated placemaking tends to obscure the complex, lived realities of communities, effectively erasing the socio-spatial diversity that originally constituted the neighborhood. This erasure indirectly facilitates gentrification by promoting a sanitised, marketable version of urban space. Real estate developers and marketers leverage social media platforms to rebrand and promote urban districts, selectively curating images and content that align with an idealised urban lifestyle-trendy cafes, boutique stores, fitness studios, and other aspirational commercial activities. These representations reinforce the upscale aesthetics of gentrifying neighborhoods, further aligning them with the tastes of affluent newcomers (Walters and Smith, 2024).

2.2 Platform User Identity and Gentrification

Digital platforms have fostered new modes of "consumption" and "leisure" that emphasize personalised experiences, cultural tastes, and identity formation. Within this framework, the middle class leverages social media to express their identity and social status, attracting like-minded groups to specific locations. This process influences neighborhood reputation, rent, and housing prices, ultimately driving gentrification. For instance, Zukin conceptualise user reviews on Yelp as a form of "discursive investing," where reviewers deploy their cultural capital through critiques, showcasing their tastes and preferences. These digital expressions attract consumer groups with similar socioeconomic backgrounds, shaping the developmental trajectory of entire neighborhoods and accelerating their restructuring and gentrification (Zukin et al., 2017). Platforms such as Yelp, TimeOut, and Eater further reshape consumers' geographical imaginations through user-generated content (reviews, descriptions, and ratings), influencing culinary and spatial aesthetics. By promoting certain local food cultures, these platforms attract new residents and investors-often at the expense of long-term residents and indigenous community culture, thereby hastening displacement and gentrification. Irene Bronsvoor's case study of Amsterdam's Javastraat district illustrates how social media platforms serve as stages for affluent groups to construct lifestyles and assert status. Through their online activities, these actors reinforce their social prestige and cultural capital, actively shaping the street's developmental direction. In doing so, they not only reflect but also amplify and reconfigure the dynamics of middle-classification, further entrenching gentrification processes (Bronsvoort and Uitermark, 2022).

2.3 Digital Platforms and the Marginalisation of Vulnerable Groups

The technological barriers inherent in digital platforms, combined with profit-driven algorithmic content distribution mechanisms, may inadvertently contribute to the marginalisation of vulnerable populations and exacerbate social segregation. This phenomenon manifests in two primary dimensions. Firstly, digital technologies can function as exclusionary mechanisms that restrict access to specific urban spaces. For instance, coworking spaces (CWS), while ostensibly designed as shared work environments, effectively operate as exclusive domains for the "mobile elite." Through financial barriers, technological requirements, and access control systems (e.g., keycard entry), these spaces systematically exclude individuals "deemed unworthy," transforming ostensibly public platforms into high-end venues catering specifically to global network elites. These spaces thus serve triple functions: as consumption sites for elites, mobility hubs for privileged professionals, and networking nodes for exclusive social circles (Fast, 2024). Secondly, the profit-oriented nature of platform capitalism leads to algorithmic mechanisms that disproportionately prioritise young, affluent demographic groups with significant consumption potential. Berglund et al.'s study of digital platforms in Detroit's downtown gentrification process reveals how platform algorithms contribute to the social isolation of elderly populations. Their research demonstrates that promotional materials on digital and social media platforms predominantly target young professionals - exemplified by initiatives like the "live downtown" campaign designed to attract this demographic to urban centers. These platform-mediated place-marketing strategies typically present sanitised, aesthetically enhanced urban imagery tailored to younger sensibilities, which may unintentionally intensify the marginalisation of vulnerable groups (particularly elderly residents) while accelerating downtown gentrification processes. This dual dynamic of technological exclusion and algorithmic bias creates a self-reinforcing cycle where platform-mediated urban development increasingly caters to privileged groups while rendering vulnerable populations increasingly invisible in both digital and physical urban spaces (Berglund et al., 2022).

In conclusion, a growing body of evidence demonstrates that user-generated content on lifestyle service platforms and social media, coupled with platform capitalism-driven discourse construction and algorithmic operations, significantly impacts neighborhood reputation, investment patterns, and demographic shifts, ultimately influencing community restructuring and gentrification processes. However, substantial research gaps remain in this field. Current discussions on digital-era gentrification predominantly remain at the level of theoretical speculation or fragmented case studies of individual communities. Existing research has primarily relied on field investigations and small-sample data from specific neighborhood cases to examine digital platforms' influence on gentrification, while neglecting macro-level analysis of how platform-generated check-in popularity and content differentially affect communities at various stages of gentrification, with a notable lack of large-scale, longitudinal data support and systematic analysis. To address these research gaps, this study selects Nanjing's main urban area as a case study. By integrating multi-source platform data from 2019 to 2024, we aim to establish a multidimensional gentrification impact factor system. Incorporating machine learning and other quantitative methods, we will investigate how check-in popularity on digital platforms influences gentrification processes, thereby providing a more comprehensive understanding of platform-mediated urban transformation mechanisms.

3.Methods

3. 1 Study Area and Data Sources

This study focuses on the main urban area of Nanjing City, Jiangsu Province, China, which comprises 600 neighborhoods (as shown in Figure 1). As a major provincial capital undergoing rapid urbanisation and digital transformation, Nanjing represents an ideal case for investigating the impact of digital platforms on gentrification processes. The research utilises two primary data sources: Beike (China's largest real estate transaction platform) and Weibo (one of China's most influential social media platforms). Specifically, we collected 37,974 property transaction records from Beike covering the period 2019-2024 within Nanjing's main urban area. These records contain detailed transaction information including neighborhood location, transaction date, average price, and floor area, as well as property characteristics such as building age, floor area ratio, and green space ratio. Concurrently, we obtained 693,638 geotagged check-in records from Weibo during the same period, each containing precise longitude and latitude coordinates, venue categories, user-generated content, engagement metrics (including reposts, comments, and likes), and user profile information (including followings, followers, gender, and IP addresses). This comprehensive, multi-source dataset provides a robust foundation for examining the mechanisms through which digital platforms influence urban gentrification.

Figure 1. The area of central Nanjing

3.2 Gentrification Measurement Methodology

In China's socioeconomic context, where residential property values serve as the primary indicator of household economic status (with real estate constituting approximately 70% of average household assets), the transaction volume of secondary housing markets provides an empirically valid measure of actual residential mobility patterns within urban neighborhoods (Shan et al., 2023). Based on this conceptual framework, our study employs property transaction data from the Beike platform to construct a multidimensional gentrification assessment system through three specifically designed indices: (1) the Average Social Stratification Index (ASSI), which quantifies neighborhood-level socioeconomic composition through property value distributions; (2) the Fixed Property Index (FPI), measuring physical asset appreciation while controlling for structural characteristics; and (3) the Value Change Index (VCI).

(1) Average Social Stratification Index (ASSI)

Based on the average transaction prices and floor areas of housing transactions in Nanjing in 2019, we first categorised the average transaction prices of properties in the main urban area of Nanjing into five tiers using the quintile method (Table 1). Through weighted summation, we obtained the community social stratification index. Given that housing constitutes the most significant household asset in China, this index can reflect the average economic-social stratification within each community. The formula is:

$$ASSI_i = \sum_{k=1}^{5} k \cdot P_{ki}$$

where k represents the weighting coefficient for each tier, with assigned values of 1 (less than 2 million CNY), 2 (2-4 million CNY), 3 (4-6 million CNY), 4 (6-8 million CNY), and 5 (greater than or equal to 8 million CNY). Pki denotes the percentage of properties in each tier within a given community.

Table 1. Property classification

Level1	Below 2 million CNY
Level 2	2-4 million CNY
Level 3	4-6 million CNY
Level 4	6-8 million CNY
Level 5	Above 8 million CNY

(2) Fixed Property Index (FPI)
The Fixed Property Index (FPI) is designed to measure whether the average household asset composition within a community exceeds Nanjing's 2019 average level. Using the mean TASSI (Total Average Social Stratification Index) of Nanjing's main urban area in 2019 as the benchmark, this index evaluates the relative asset level of communities within the city's central districts.The calculation formula is:

$$FPI_i^{2019} = \frac{ASSI_i^{2019}}{TASSI^{2019}}$$

Where:ASSI_i2019 represents the ASSI value of individual communities in 2019;TASSI_2019 denotes the aggregate ASSI value of all communities in Nanjing during 2019

(3) Value Change Index (VCI)
The Value Change Index (VCI) measures the transformation in household asset composition across communities between 2019 and 2024. By comparing the relative change in community-level ASSI to the citywide TASSI during this period, this index isolates gentrification-driven social replacement from general market fluctuations. The formula is expressed as:

$$VCI_i = (ASSI_{i2024}/ASI_{i2019})/(TASSI_{2024}/TASSI_{2019})$$

Where:
ASSI_i2019 and ASSI_i2024 represent the ASSI values for individual communities in 2019 and 2024 respectively.
TASSI_2019 and TASSI_2024 denote the aggregate ASSI values for all Nanjing communities in the corresponding years.

Finally, the gentrification stages of communities in Nanjing's main urban area were determined based on their FPI and VCI values, with specific classification criteria as shown in Table 2.

Table 2. Gentrification Stage Classification Criteria

Ungentrified	FPI2019< 1 and VCI < 1, set Gentri-type = 0;
Gentrifying	FPI2019< 1 and VCI > 1, set Gentri-type = 1;
Gentrified	FPI2019>1 and VCI < 1, set Gentri-type = 2;
Re-gentrifying	FPI2019> 1 and VCI > 1, set Gentri-type = 3;

3.3 Selection of Influencing Factors
Digital footprints have been empirically demonstrated to provide early-warning signals for urban gentrification processes. The "check-in" popularity on social media platforms, visitor demographics, and spatial distribution patterns exhibit strong coupling with neighborhood value reassessment and consumption upgrading (Chapple et al., 2022). In the Chinese context, Weibo's advantages - including its massive user base, real-time information flow, and geotagging features - have made it widely adopted for studies on urban vitality and public space utilisation. This study systematically constructs Weibo check-in independent variables across three dimensions:

check-in popularity, user demographics, and location characteristics.

Table 3. Independent Variables

Dimension	Variable ID	Variable Name	Mean
Neighborhood Check-in Popularity	X1	Total number of Check-ins	1170.15
	X2	Total interactions of Check-ins	14861.55
	X3	Total reposts of Check-ins	1592.56
	X4	Total comments of Check-ins	3157.81
	X5	Total likes of Check-ins	10111.19
Neighborhood Check-in User Demographics	X6	Normal user check-ins	927.50
	X7	KOC check-ins	203.47
	X8	KOL check-ins	39.18
	X9	Female check-ins	711.21
	X10	Male check-ins	458.94
	X11	Nanjing resident check-ins	398.74
	X12	Non-Nanjing resident check-ins	771.41
Neighborhood Check-in Locations	X13	Check-ins at residential locations	266.11
	X14	Check-ins at commercial locations	341.02
	X15	Check-ins at workplace locations	93.27
	X16	Check-ins at university locations	77.71
	X17	Check-ins at public service locations	62.67
	X18	Check-ins at parks	59.08
	X19	Check-ins at tourist attractions	95.21
	X20	Check-ins at cultural & sports venues	58.80
	X21	Check-ins at public transportation facilities	45.99
	X22	Check-ins at other locations	41.145

3.4 Local Bivariate Moran's I

To characterise the spatial coupling between Weibo check-in activity and different gentrification stages at micro-scale levels, this study employs the local bivariate Moran's index to detect spatial associations between these variables. The computational formula is:

$$I_i^{XY} = z_{x_i} \sum_{j=1}^{n} w_{ij} z_{y_j},$$

Where:

$z_{xi} = (x_i - x\text{-})/s_x$ and $z_{yj} = (y_j - Y)/s_y$ represent standardised residuals of gentrification stage values and Weibo check-in counts respectively; w_{ij} denotes row-standardised spatial weight elements.

3.5 Random Forest Model

Given the multiclass categorical nature of the dependent variable, this study employs a Random Forest classification model (Breiman, 2001) as the primary analytical framework. As an ensemble tree-based learning algorithm, the Random Forest classifier comprises multiple decision trees, each constructed through bootstrap sampling of observations and random feature selection from the dataset. This non-parametric model demonstrates particular strengths in handling high-dimensional, multiclass data with complex nonlinear relationships, while exhibiting superior generalisation capacity and robustness against overfitting - characteristics that have led to its increasing adoption in urban studies(Deb and Smith, 2021; Kim, 2022). The model was implemented using Python's scikit-learn library, with the dataset partitioned into training and testing subsets at an 80:20 ratio. Model performance was evaluated using standard classification metrics including accuracy and F1-score on the test set. To ensure model stability and reproducibility, hyperparameter tuning and validation were conducted through 5-fold cross-validation.

3.6 Shapley Additive Explanations (SHAP)

To elucidate the operational mechanisms and relative importance of predictive factors in gentrification classification, this study incorporates SHAP (SHapley Additive exPlanations) for both global and local interpretation of the trained Random Forest model (Lundberg and Lee, 2017). Grounded in cooperative game theory and Shapley value principles, SHAP provides a theoretically-grounded framework for quantifying the marginal contribution of each feature to model predictions. This approach not only yields feature importance rankings but also reveals the directional effects (positive/negative) of specific feature value ranges on classification outcomes, thereby effectively demystifying the "black box" nature of Random Forest models and enhancing overall interpretability. All modeling and visualisation procedures were implemented in Python 3.7, utilizing established open-source libraries including scikit-learn for machine learning, pandas for data manipulation, matplotlib for visualisation, and the specialised shap package for model interpretation.

4.Results

4.1 Spatial Distribution of Gentrified Communities in Nanjing

The main urban area of Nanjing comprises 600 communities, which were classified into four categories based on our indicator calculations: 107 non-gentrified communities, 246 communities undergoing gentrification, 155 gentrified communities, and 92 re-gentrified communities. The results demonstrate that most regenerated communities in Nanjing exhibit varying degrees of gentrification. The gentrified and re-gentrified communities are predominantly concentrated in several key areas: Jianye District, the old urban areas of Gulou and Qinhuai Districts, the riverside areas of Pukou District, the Southern New City of Jiangning District, and the eastern Zijin Mountain area of Qixia District. These areas correspond precisely with recent government-led urban redevelopment projects, including the urban renewal initiatives in Gulou District and Qinhuai old town, as well as new urban developments such as the Hexi New City in Jianye District, the core area of Jiangbei New District, Southern New City, and Zidong New District.

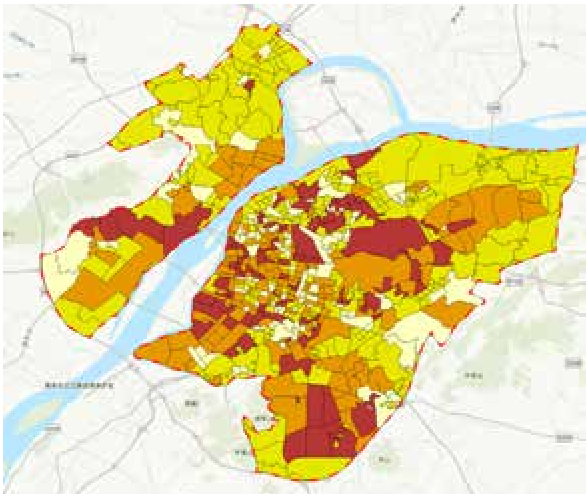


Figure 2: Gentrification Stages in Central Nanjing

4.2 Local Bivariate Moran's Index Analysis

The research findings reveal a significant spatial correlation between the gentrification index of communities in Nanjing's main urban area and the popularity of check-in activities on social platforms. Distinct spatial clustering patterns emerge: "high-high" agglomeration is observed in communities surrounding commercial centers, central business districts, tourist attractions, and university towns, while "low-low" agglomeration characterises peripheral urban areas. This spatial association demonstrates the close relationship between digital platform activities and urban gentrification processes, particularly in areas undergoing rapid urban transformation and redevelopment.

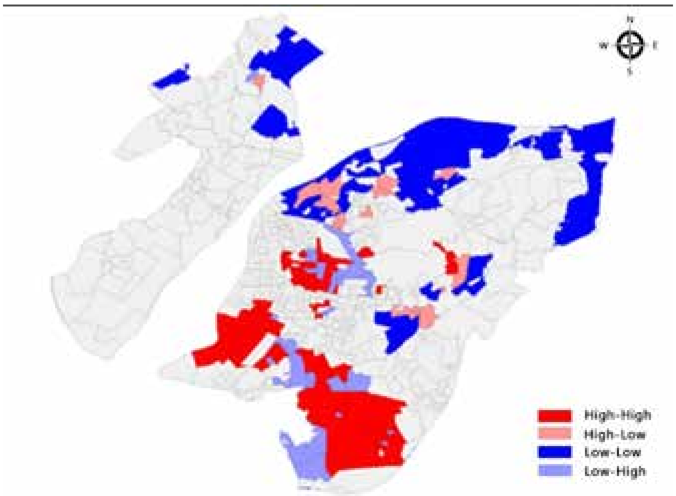


Figure 3: Gentrification Stages in Central Nanjing

4.3 Regression Results

To comprehensively evaluate the model's discriminative capability for identifying community gentrification stages, this study employed a dual analytical approach. First, we constructed Receiver Operating Characteristic (ROC) curves (Figure 4) and calculated the Area Under Curve (AUC) values for each classification category. The results demonstrate an average AUC≈0.68, indicating that the model exhibits reasonable feasibility in distinguishing communities at different gentrification stages.

For deeper insight into the key variables influencing gentrification stage evolution, we computed SHAP values from the trained multiclass Random Forest model and generated both global (Figure 5) and class-specific feature importance Pareto charts (Figures 6a-d). These visualisations present the hierarchical ranking and cumulative contribution of variables across all samples and for each of the four gentrification types, with detailed cumulative contribution rates provided in Table 4.

The analysis reveals three key findings regarding variable contributions: (1) Overall community check-in popularity metrics (X1-X5) demonstrate the highest explanatory power, accounting for 39.58% of cumulative contribution; (2) Neighborhood check-in user demographics (X6-X12) show moderate influence at 35.8%; while (3) Neighborhood check-in location characteristics (X13-X22) exhibit relatively minimal impact at 24.62%. This pattern underscores the predominant role of community check-in popularity in gentrification processes, particularly for communities in the ongoing and completed gentrification stages. Conversely, the specific types of check-in locations appear to exert comparatively weaker effects on gentrification dynamics.

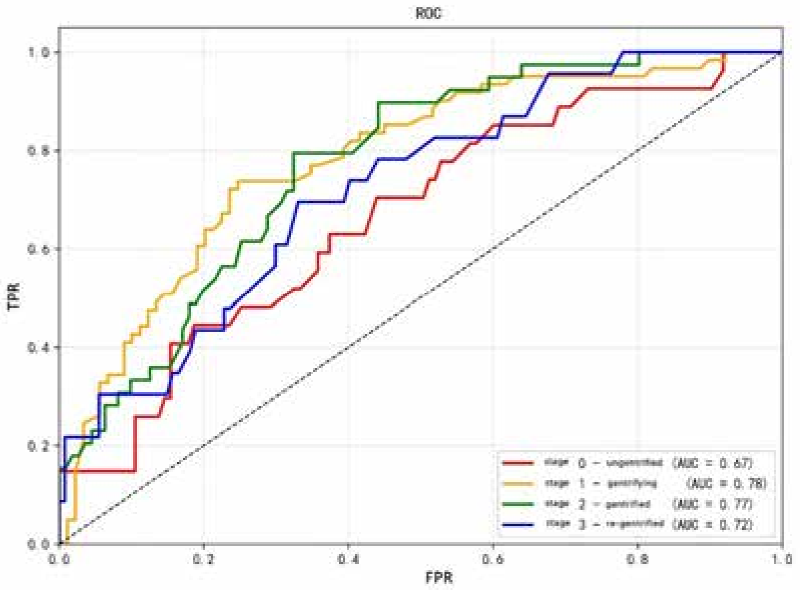


Figure 4. Multiclass ROC Curve

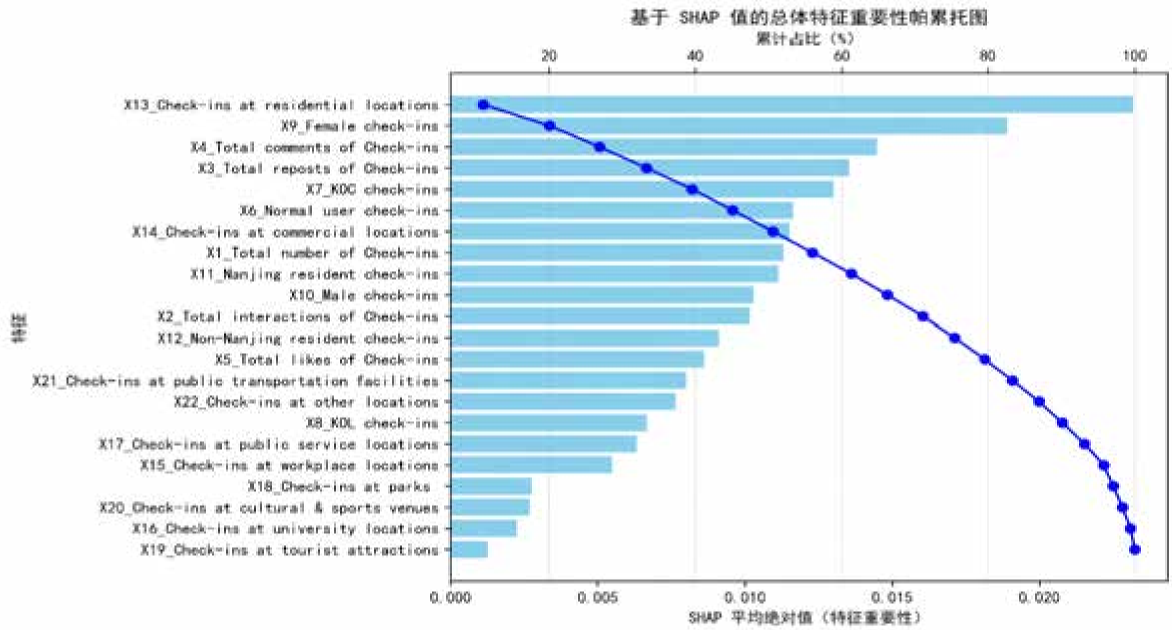


Figure 5. Pareto chart of feature importance

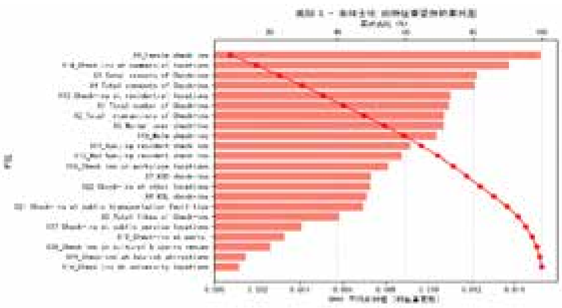


Figure 6-a. Pareto chart of feature importance for ungentrified stage

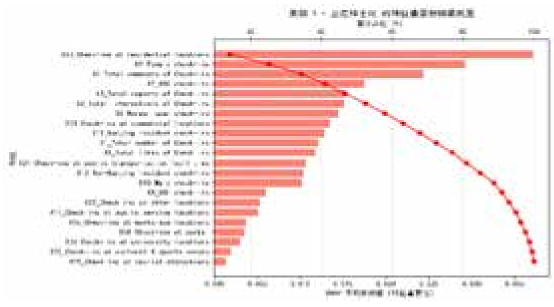


Figure 6-b. Pareto chart of feature importance for gentrifying stage

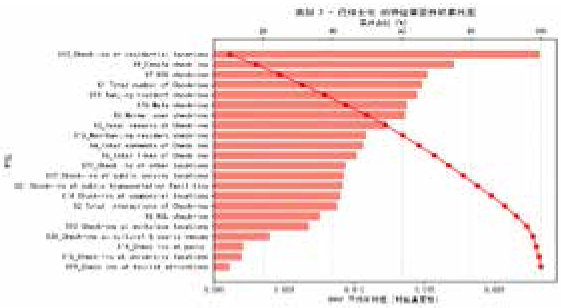


Figure 6-c. Pareto chart of feature importance for gentrified stage

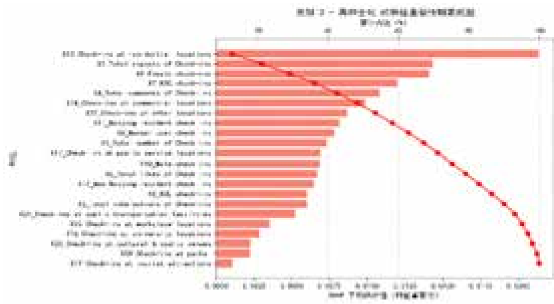


Figure 6-c. Pareto chart of feature importance for re-gentrifying stage

Table 4. Precision and Recall Values of the Random Forest Model

Stage of Gentrification	Neighborhood Check-in Popularity (X1–X5)	Neighborhood Check-in User Demographics (X6–X12)	Neighborhood Check-in Locations (X13–X22)
Average	39.58%	35.80%	24.62%
Ungentrified	35.79%	38.13%	26.08%
Gentrifying	46.72%	34.22%	19.06%
Gentrified	37.68%	35.82%	26.50%
Re-gentrifying	42.54%	32.65%	24.81%

To elucidate the specific mechanisms through which different variables influence model outputs, this study conducted interpretability analysis using SHAP value distribution plots (Beeswarm Plots) for the classification models of four gentrification stages, with results presented in Figures 6a-d. Each plot displays the SHAP value distribution of variables for the respective category, where the horizontal axis represents the marginal effect of each variable on model output, and the color gradient indicates the magnitude of variable values. Comparative analysis reveals significant differences in variable effect mechanisms across distinct gentrification stages.

The analysis reveals distinct patterns in the marginal effects of Weibo check-ins across different gentrification stages. Notably, check-in activities demonstrate predominantly negative marginal effects in non-gentrified and ongoing gentrification areas, while exhibiting significantly positive effects in gentrified and re-gentrified communities.

In ungentrified and gentrifying neighborhoods, model predictions primarily rely on variables including female user check-ins (X9), commercial location check-ins (X14), interaction metrics (e.g., X3 reposts and X4 comments), and residential area check-ins (X13). The SHAP values for these variables predominantly cluster near zero, indicating relatively moderate influences on model outputs. Nevertheless, these patterns suggest that lower female activity and weaker commercial engagement represent key characteristics of non-gentrified areas. Overall, these communities exhibit distinctive features of limited platform engagement, low social interactivity, and predominantly residential space utilisation.

For gentrified areas, the discriminative factors concentrate on residential area check-ins (X13), female user check-ins (X9), Key Opinion Consumer (KOC) check-ins (X7), and total check-in volume (X1). These variables reflect high-frequency, targeted social activity patterns, indicating these areas have developed mature "internet-famous" attributes. Additionally, check-ins by Nanjing local residents (X11), male users (X10), and content dissemination metrics like reposts and comments demonstrate significant influences, suggesting synergistic effects between online reputation and local participation in reinforcing community stability.

Re-gentrified areas display complex, multi-source participation characteristics. Decision-making incorporates not only traditional factors like residential check-ins (X13), reposts (X3), and female check-ins (X9), but also variables such as KOC check-ins (X7), comments (X4), and commercial location check-ins (X14). This broad participation spectrum indicates that after one or multiple rounds of redevelopment, these areas have evolved into urban spaces characterised by diverse user coexistence and complex dissemination mechanisms. The more dispersed SHAP values and varied effect directions reveal higher dynamism and diversity in re-gentrified areas, potentially featuring composite social nodes that combine both consumption power and community cohesion.

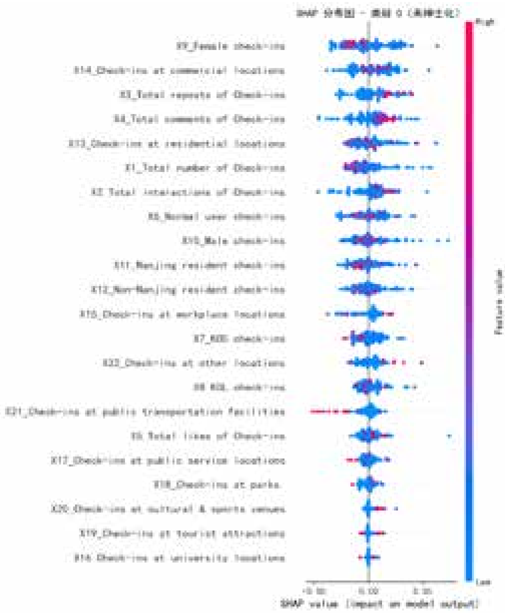


Figure 7-a. beeswarm chart for ungentrified stage

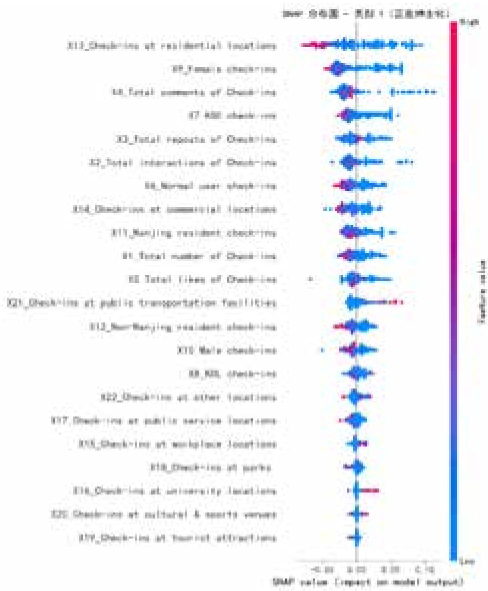
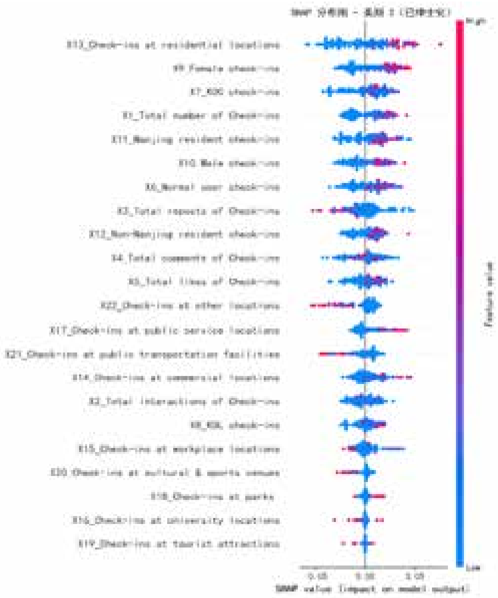


Figure 7-b. beeswarm chart for gentrifying stage



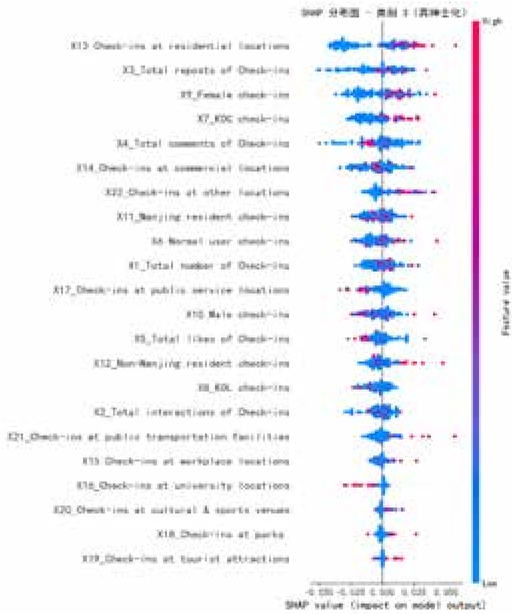


Figure 7-c. beeswarm chart for gentrified stage

Figure 7-d. beeswarm chart for re-gentrified stage

5.Conclusion

This study pioneers an analytical approach that integrates multi-platform data from Beike and Weibo with interpretable machine learning models, establishing a behavioral-oriented research paradigm for understanding urban gentrification in the digital era. The findings yield significant theoretical contributions by demonstrating that social media engagement patterns - particularly check-in frequency and interaction levels - exhibit strong correlations with neighborhood gentrification processes. Notably, the activities of female users, Key Opinion Consumers (KOCs), and local residents exert disproportionately greater influence on gentrification dynamics compared to other demographic groups. The analysis further reveals that gentrification stages show the strongest association with check-in intensity in residential compounds and commercial districts.

The study elucidates how digital capital transcends spatial-temporal constraints to commodify urban space through platform-mediated mechanisms. Algorithmic governance structures guide social media users to anchor textual and visual content to specific physical locations through geotagging practices. This process of 'discursive investing' systematically attracts socioeconomically homogeneous groups and investors to targeted neighborhoods, thereby accelerating gentrification through subtle, digitally-mediated pathways. These insights necessitate critical policy considerations: urban governance frameworks must incorporate the impacts of high-engagement digital activities on community transformation, particularly regarding demographic restructuring and emerging gentrification landscapes. Concurrently, regulatory interventions should address platform algorithms and content curation to mitigate inter-neighborhood disparities and advance spatial justice in platform urbanism.

While this research provides macro-level evidence through the analysis of Weibo check-in data in Nanjing, the complex mechanisms of platform-driven gentrification warrant further investigation. Future studies should incorporate multi-dimensional data streams (including textual comments, visual imagery, and mobility traces) from representative urban contexts to more comprehensively capture the logic of digitally-mediated spatial reproduction. Particularly crucial is examining the symbiotic relationship between algorithmic recommendation systems and commercial capital accumulation in shaping gentrification outcomes. Such inquiries would strengthen the theoretical foundations for platform governance and equitable urban policy frameworks in the digital age.

Reference

Berglund, Lisa, et al (2022) MISSING OLDER ADULTS IN A GENTRIFYING DOWNTOWN: Detroit's Rebrand for a Young and Talented Pool of Residents. International Journal of Urban and Regional Research, 46(6), pp.973-997.

Breiman, Leo (2001) Random forests. Machine learning, 45, pp.5-32.

Bronsvoort, Irene and Uitermark, Justus L (2022) Seeing the street through Instagram. Digital platforms and the amplification of gentrification. Urban Studies, 59(14), pp. 2857-2874.

Chapple, Karen, et al (2022) Monitoring streets through tweets: Using user-generated geographic information to predict gentrification and displacement. Environment and Planning B: Urban Analytics and City Science, 49(2), pp.704-721.

Deb, Debzani and Russell M, Smith (2021). Application of random forest and SHAP tree explainer in exploring spatial (in) justice to aid urban planning. ISPRS International Journal of Geo-Information, 10(9), pp.629.

Fast, Karin (2024) Who has the right to the coworking space? Reframing platformed workspaces as elite territory in the geomedia city. Space and Culture, 27(1), pp.48-62.

Glass, Ruth (1964) Aspects of Change. London: MacGibbon & Kee

Hartmann, Maren, and André Jansson (2024) Gentrification and the right to the geomedia city. Space and Culture, 27(1), pp.4-13.

Kim, Yesuel, and Youngchul Kim (2022) Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models. Sustainable Cities and Society, 79, pp.103677.

Lees, Loretta, Hyun Bang Shin, and Ernesto López-Morales (2016) Planetary gentrification. John Wiley & Sons.

Lundberg, Scott M., and Su-In Lee (2017) A Unified Approach to Interpreting Model Predictions. Advances in neural information processing systems, 30.

Polson, Erika (2024) From the tag to the# hashtag: Street art, Instagram, and gentrification. Space and Culture, 27(1), pp.79-93.

Shan, Lu, He, Shenjing and Wan, Chenxin (2023). Unraveling the dynamic Airbnb-gentrification interrelation before and after the COVID-19 Pandemic: Evidence from Beijing, China. Cities, 137, pp.104270.

Törnberg, Petter, and Uitermark, Justus (2022) Urban mediatization and planetary gentrification: The rise and fall of a Favela across media platforms. City and Community, 21(4), pp. 340-361.

Walters, Peter, and Smith, Naomi (2024) It's so ridiculously soulless: Geolocative media, place and third wave gentrification. Space and Culture, 27(1), pp. 94-109.

Zukin, Sharon, Scarlett Lindeman, and Laurie Hurson (2017) The omnivore's neighborhood? Online restaurant reviews, race, and gentrification. Journal of Consumer Culture, 17(3), pp.459-479.