Sustainable Tourism through Locality-Aware Recommendations

Furkan Çelik[1,*], Ashmi Banerjee[1] and Wolfgang Wörndl[1]

[1] *Technical University of Munich, Germany*

Abstract

Urban point-of-interest (POI) recommenders influence the mobility patterns of tourists and residents, thereby affecting which venues receive visitors. These systems commonly rely on popularity and proximity, which can cause densely populated hotspots while under-representing locally oriented venues in recommendation lists. This imbalance raises sustainability concerns for urban tourism, as concentration in a small set of hotspots can contribute to overtourism while other neighborhoods remain largely underexplored. For this reason, we propose a continuous, city-relative locality feature per venue that captures how strongly a venue is associated with users whose activity is concentrated in the target city or users whose activity is more geographically dispersed. This feature is derived from user review distributions in the Yelp Open Dataset for restaurants in Philadelphia. Furthermore, we convert this feature into a score and integrate it as a tunable signal into a transparent ranker and compare locality-aware variants against a baseline using both accuracy and sustainability-oriented metrics. Our contributions include data-driven mapping of restaurants along a continuous locality spectrum, integrating this locality score into a standard POI ranking, and analyzing its impact on relevance, novelty, geographic spread, and catalog coverage. Our results show that integrating locality can improve sustainability-oriented objectives such as novelty and coverage, and shift recommendations toward more locally oriented restaurants.

Keywords

Tourism, Recommender Systems, Sustainability, Locality, Novelty

1. Introduction

Tourism encompasses a wide range of activities, from visiting iconic attractions to immersing oneself in everyday local life. While some visitors follow well-established tourist itineraries, others seek experiences that are more in line with residents' daily routines. These activities correspond to distinct mobility patterns. One standard grouping of mobility patterns separates them into tourists and residents [1]; however, building on digital footprints [2], these groupings can be further expanded [3]. Common to all these patterns is participation in everyday urban life through the shared use of city infrastructure. This type of tourism is researched under the domain of urban tourism [4], which is a significant driving force behind the urban socio-spatial change, making the investigation of different mobility patterns important and motivating in this domain.

In many popular urban destinations, tourist spaces are increasingly interwoven with residents' everyday living spaces, creating socio-spatial interactions that have important implications for urban planning and tourism management [5]. These interactions must be taken into consideration when promoting or highlighting specific regions; otherwise, visitors from any group may cluster in a small set of highly promoted areas, while other neighborhoods are visited only by some locals. One example of such clusters is often discussed through the concept of the tourist bubble: regions planned and serviced primarily for tourists, rich in attractions and short-term accommodation, and connected by recognizable paths between "must see" sights [6]. While tourist bubbles can provide convenience and perceived safety, they may reduce contact with everyday local life and increase vulnerability to risks concentrated in central areas. One such risk is overtourism, where excessive concentration can harm residents' quality of life and degrade visitor experience through crowding, noise, and rising costs [7, 8].

Joint Proceedings of the ACM UMAP Workshops 2026, UMAP 2026, June 8–11, 2026, Gothenburg, Sweden
*Corresponding author.
f.celik@tum.de (F. Çelik); ashmi.banerjee@tum.de (A. Banerjee); woerndl@in.tum.de (W. Wörndl)
© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-4206/UMAP4GOOD-5.pdf
CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-05-28

On the other hand, many urban regions are primarily visited by residents, forming another cluster of spaces shaped by daily routines and social life that remain largely underpromoted. Prior work using geotagged social media data shows that tourist hotspots tend to be more centralized and spatially concentrated, as opposed to the dispersed mobility of locals across the city, indicating a measurable separation between visitor and resident spaces [5, 9]. Recognizing these spatial patterns is crucial for both urban management and digital platforms that influence users' mobility decisions. In particular, these insights can support strategies to reduce congestion in hotspots by distributing visitor flows more evenly across the urban space [10].

Recommender systems (RS) have become essential tools in this context. Tourism recommender systems (TRS) help users navigate large volumes of information and receive personalized suggestions for destinations, attractions, and services [11]. In urban settings, location-based point-of-interest (POI) recommenders are a core component of TRS. Many TRS approaches favor already popular and heavily visited venues, creating a loop that leads to overtourism unless this bias towards popularity is explicitly monitored and mitigated [12]. As a consequence, a small set of venues may receive disproportionate visibility in the recommendation lists, while many potentially suitable local venues remain unlikely to appear.

At the same time, many tourists increasingly seek to experience local culture and everyday life beyond these famous and popular places [13]. To bridge the gap between the demand for local experiences and the recommendation of such venues in TRS, we propose a continuous, city-relative *locality score* per venue. Focusing on restaurants in Philadelphia (PA, USA) from the Yelp Open Dataset, we derive a locality feature from user review distributions that indicates how strongly each restaurant is associated with users whose activity is concentrated in the city compared to users whose activity is more geographically dispersed. We then transform this feature into an interpretable locality score and integrate it into a baseline ranker.

Rather than competing with state-of-the-art deep learning models on predictive accuracy, our goal is to investigate whether this locality signal can serve as a practical control knob to steer recommendations toward more locally oriented venues and to promote sustainability-oriented objectives while maintaining acceptable ranking relevance. These sustainability objectives include novelty [14], catalog coverage [12] and geographic spread, which reflect how broadly recommendations are spread across venues in recommendation lists.

The main contributions of this paper are:

- We introduce a venue-level locality estimation procedure that places venues on a continuous spectrum ranging from tourist-oriented to locally oriented, using behavioral and city-relative signals.
- We integrate the resulting locality score into a classical ranker through a tunable parameter, enabling controllable trade-offs between ranking utility and sustainability-oriented objectives.
- We perform an evaluation on Yelp data for Philadelphia, quantifying how varying the locality weight reshapes recommendation lists in terms of relevance, novelty, spatial dispersion, item coverage, and the locality orientation of recommended venues.

The remainder of the paper is organized as follows. Section 2 reviews related work on tourism recommender systems, urban mobility, and locality-aware recommendation. Section 3 formalizes the locality score construction and the ranking models. Section 4 presents the experimental setup, including the dataset, configuration grid, evaluation metrics, and robustness checks, followed by the results and discussion in Section 5. Section 6 concludes the paper, outlines limitations, and directions for future work.

2. Background and Related Work

Tourism recommender systems (TRS) adapt traditional recommendation methods to a domain where interactions are short-lived, user intent and context change rapidly, and the item space is heterogeneous (e.g., venues, attractions, routes, and activities) [15]. User behavior in this domain is typically characterized as either session-based or trip-based, rather than over a long period. These properties reduce available historical information and increase the value of models that can learn from short sequences and immediate context [16]. In addition, the focus on urban regions adds geographic constraints and dense spatial competition among the recommended items, making transparency and controllability in ranking especially relevant for different stakeholders. These characteristics motivate designs that combine simple yet robust signals (e.g., popularity, distance) with interpretable domain features that reflect tourist and local dynamics.

Accordingly, many TRS adopt context-aware recommendation frameworks to operationalize these constraints. Context-aware recommenders can model additional information, such as time, feasibility of user choices, mobility, and user location, to make personalized recommendations [17]. In this context, a key signal is the distance between the user's request location and the candidate recommendations. The probability of visiting a POI typically decreases as the distance from the user's request location increases [18]. As a result, ranking POIs by a combination of popularity and proximity has become one of the standard baseline approaches, while leaving gaps to incorporate additional, interpretable signals, such as locality. In POI recommendation, user–venue interactions may include repeated engagements with the same place, are highly sparse, and exhibit pronounced spatiotemporal regularities (e.g., distance decay, seasonal cycles) [18, 19]. Furthermore, many mobility studies reveal regularities in urban mobility and exploration, identifying intra-city hotspots that result in localized activity peaks and uneven spatial coverage across neighborhoods [20, 21, 22].

Location-based social network (LBSN) applications, including Foursquare and Yelp, provide information that is widely used for POI recommendations. These platforms enable users to log check-ins or submit reviews across cities worldwide. Many studies treat each city or region as an independent dataset for recommendation [23, 24], which is both practical and reasonable given that user interest is typically focused on venues within a target destination rather than across cities.

Evaluation in these systems often follows top-k offline settings, with metrics such as precision and recall, using historical interactions to approximate future recommendations more accurately [19]. However, accuracy metrics alone do not account for dense population in specific regions, catalog coverage, or user experience goals such as discovery and variety. These ranked lists also influence the visibility of venues and can amplify concentration in certain regions. Patro et al. [25] demonstrate that conventional top-k recommendations on Yelp and Google Local can concentrate attention on a small subset of businesses, risking overcrowding at some locations while leaving others with lower visitations. They further report that naive attention balancing strategies (e.g., poorest-k) can reduce concentration but at a significant utility cost (about 60-70%), motivating multi-objective evaluation beyond accuracy. More recently, Banerjee et al. [26] proposed a composite sustainability indicator for city trip recommenders that combines CO_2e emissions, destination popularity, and seasonality, demonstrating how interpretable sustainability signals can be integrated into tourism recommender systems.

Locality as a ranking signal is directly relevant in this context. Zhao et al. [27] define locally interesting venues as places that locals frequently visit but are obscure to most foreigners, and leverage LBSN check-in traces to recommend such venues to tourists. Their framework extracts users' latent social dimensions via Bayesian probabilistic tensor factorization with social and location regularization, detects and profiles local interest communities, and performs cross-region community matching to transfer local preferences to foreign-city recommendations. They report that identifying locally interesting venues can improve recommendation accuracy. However, their framework operationalizes locally interesting venues as a discrete category based on local and foreign visit proportions over check-in data, while the recommendation model itself relies on latent tensor factorization, which limits interpretability. Moreover, their evaluation focuses primarily on accuracy metrics rather than sustainability-oriented objectives. In contrast, we propose a continuous, city-relative locality score derived from review activity and evaluate it alongside both accuracy and sustainability-oriented objectives, demonstrating how this signal can promote sustainable urban tourism within a transparent ranking framework.

3. Methodology

In this section, we formally define the methodology. While the method is designed for review-based LBSNs, the general approach could potentially be adapted to other platforms with similar interaction data. Throughout the paper, we use "locally oriented" and "tourist-oriented" as shorthand for the two directions of a continuous, city-relative locality spectrum; these terms do not denote ground-truth user or venue classes.

3.1. Problem Definition

Let U denote users, I businesses, and C cities. Each business $i \in I$ has coordinates $\mathbf{x}_i \in \mathbb{R}^2$, city $c(i) \in C$, and categories $\mathrm{cat}(i)$. Interactions are observed as timestamped reviews

$$\mathcal{D} = \{(u, i, t) : u \in \mathcal{U},\ i \in \mathcal{I},\ t \in \mathbb{T}\},$$

where t is the review time.

We limit the restaurant catalog of the city c to the set of open venues labeled as restaurants, denoted by $I_c = \{\, i \in I : c(i) = c,\ \texttt{Restaurants} \in \mathrm{cat}(i),\ \texttt{is\_open} = 1 \,\}$. Given a user $u \in U$, we aim to recommend a ranked top-k list $\pi_{u,c}^{(k)}$ from this catalog that favors locally oriented venues.

To promote discovery, we exclude from a user's candidate set the restaurants they interacted with during the training period:

$$I_{u,c} = I_c \setminus I_u^{\mathrm{tr}},$$

where I_u^{tr} is the set of items reviewed by u in the training split (within c). Recommendations are produced as

$$\pi_{u,c}^{(k)} = \mathrm{TopK}_{i \in I_{u,c}}\ s(i \mid u, c),$$

where $s(i \mid u, c)$ is a scoring function that combines interpretable signals (popularity, proximity) with a venue-level locality feature.

Query location and distance. For each test user, we define a query location $\mathbf{q}_u$ as the geographic centroid of the user's training-period review coordinates in city c. For a candidate restaurant i, we compute the geographic distance $d(i, \mathbf{q}_u)$ with Haversine [28] on latitude/longitude, and use proximity as a decreasing function of distance (closer restaurants receive higher proximity scores). For cold-start users with no training reviews in city c, $\mathbf{q}_u$ is undefined and the proximity feature is omitted.

3.2. Deriving a Continuous Locality Score

Our key contribution is an interpretable, continuous, venue-level locality signal computed from review activity, avoiding a discrete tourist/local classification. The derivation proceeds in two steps: we first compute a user-level locality ratio $\ell_{u,c}$ capturing how concentrated each user's reviewing activity is in city c, then aggregate these ratios across a restaurant's reviewers to obtain a venue-level locality proportion p_i that reflects the locality profile of its reviewer base.

User–city locality ratio. Following behavioral approaches to distinguishing local and tourist activity patterns in LBSNs [1, 2, 3], we define the user–city locality ratio as a continuous measure of how concentrated a user's reviewing activity is within a given city. For each user u and city c, let $n_{u,c}^{\mathrm{tr}}$ be the number of the user's training reviews in c and n_u^{tr} the user's total training reviews across all cities:

$$\ell_{u,c} = \frac{n_{u,c}^{\mathrm{tr}}}{n_u^{\mathrm{tr}}} \in [0, 1],$$

which equals 1 for users whose training activity is entirely in city c and decreases as mobility across cities increases.

Restaurant locality proportion. Similarly to prior work [27], which characterizes locally interesting venues by the proportion of local versus foreign visitors, we generalize this idea to a continuous venue-level signal by averaging the locality ratios of all reviewers of restaurant i in the training split, after deduplicating to one interaction per user–business pair. Let U_i be the set of unique users who reviewed restaurant i in the training split:

$$p_i = \frac{1}{m_i} \sum_{u \in U_i} \ell_{u,c}, \qquad m_i = |U_i|.$$

Empirical-Bayes shrinkage toward the city norm. The naive restaurant estimate p_i can be unreliable for venues with few unique reviewers. With small m_i, a handful of highly local users can inflate p_i, yielding extreme values that are not comparable to estimates supported by larger reviewer sets. In practice, we observed that this small sample effect causes many restaurants to appear similar and reducing separability. Therefore, we shrink restaurant estimates toward a city-wide baseline p_{city} using an Empirical-Bayes (EB) [29] shrinkage estimator to obtain more robust estimates and improve discrimination.

We compute the city baseline as the mean locality proportion across restaurants in the city catalog,

$$p_{\text{city}} = \frac{1}{|I_c|} \sum_{i \in I_c} p_i,$$

where each venue contributes equally to the prior, preventing high-traffic venues from dominating the city baseline. We define the shrinkage-adjusted restaurant locality as

$$\tilde{p}_i = \frac{m_i p_i + s\, p_{\text{city}}}{m_i + s},$$

where $s > 0$ controls the strength of the prior. We interpret s as a prior sample size and p_{city} as a computed prior mean from the training catalog; s is treated as a tuning parameter rather than fitted by likelihood, as the robustness checks show the qualitative tradeoff is stable across reasonable values of s. This approach downweights noisy estimates for small m_i while leaving well-reviewed restaurants largely unchanged. For restaurants with no reviewers in training ($m_i = 0$), $\tilde{p}_i$ defaults to p_{city}.

Monotone transform. The logit function is a standard monotone bijection from $(0, 1)$ to $\mathbb{R}$, widely used to linearize probability scales in statistical modeling [29]. We convert the restaurant locality $\tilde{p}_i$ into an unbounded, signed deviation from the city baseline using a log-odds offset. Concretely, after clipping probabilities to $[\varepsilon, 1-\varepsilon]$ (here $\varepsilon = 10^{-6}$), we define

$$\Delta_i \;=\; \operatorname{logit}(\tilde{p}_i) - \operatorname{logit}(p_{\text{city}}), \qquad \operatorname{logit}(x) = \log \frac{x}{1-x}.$$

Positive values indicate a stronger local orientation than the city average, and negative values indicate a more touristic orientation. Finally, we standardize Δ_i within each city to obtain the locality feature $z_{\text{loc}}(i, c)$ used by our ranking models; restaurants with $m_i = 0$ receive a neutral value $z_{\text{loc}}(i, c) = 0$.

Correlation check against popularity. To ensure that locality does not strongly correlate with popularity, we compute the association between the venue locality feature $z_{\text{loc}}(i, c)$ and a popularity proxy $\operatorname{pop}(i)$ on the training split using both Pearson (linear) and Spearman (rank) correlation tests, following prior work [30] that uses correlation analysis to validate newly proposed indices against existing indicators.

3.3. Feature Definitions

For ranking the recommendations, we use three features: popularity, proximity, and locality. To ensure comparability across features, all features are z-score standardized using statistics computed on the training split.

Popularity. Let $\texttt{reviewer\_count}_i^{\text{tr}}$ be the number of *unique* reviewers of restaurant i in the training split. We apply a logarithmic transform to map the reviewer count to a continuous scale, making it compatible with the other features:

$$\text{pop}(i) = \log\bigl(1 + \texttt{reviewer\_count}_i^{\text{tr}}\bigr),$$

and standardize it per city to obtain $z_{\text{pop}}(i)$.

Proximity. Geographic proximity is a strong predictor of POI visits, as users tend to visit venues near to their location [18]. Given query location $\mathbf{q}_u$, we compute the Haversine distance $d(i, \mathbf{q}_u)$ and set $\text{prox}(i, u) = -d(i, \mathbf{q}_u)$ (larger is closer). Proximity is standardized within the user's recommendation list to obtain $z_{\text{prox}}(i, u)$. If the per-user standard deviation is near zero, we set $z_{\text{prox}}(\cdot, u) \equiv 0$ to avoid numerical noise.

Locality. Locality is a continuous feature that captures how strongly a venue is associated with locally concentrated user activity per venue derived in Section 3.2, relative to the city average.

3.4. Scoring Rankers

Ranking POIs by a weighted combination of popularity and proximity is a well-established baseline approach in location-aware recommendation [18, 12]. We adopt this as our baseline ranker and extend it by integrating the locality feature as a third signal with a tunable weight. Let $z_{\text{pop}}(i)$, $z_{\text{prox}}(i, u)$, and $z_{\text{loc}}(i, c)$ be the standardized features. We employ simple linear scoring functions to generate a top-k recommendation list for each user.

Baseline ranker. Popularity–proximity:

$$s_{\text{B1}}(i \mid u, c; \beta) = \beta\, z_{\text{pop}}(i) + (1 - \beta)\, z_{\text{prox}}(i, u), \quad \beta \in [0, 1].$$

Locality-integrated ranker. We add locality with weight $\alpha \in [0, 1]$:

$$s_{\text{B2}}(i \mid u, c; \alpha, \beta) = (1 - \alpha)\Bigl[\beta\, z_{\text{pop}}(i) + (1 - \beta)\, z_{\text{prox}}(i, u)\Bigr] + \alpha\, z_{\text{loc}}(i, c).$$

Equivalently, s_{B2} is a convex combination of the three standardized components with weights $\bigl((1 - \alpha)\beta,\ (1 - \alpha)(1 - \beta),\ \alpha\bigr)$.

Cold-start handling. If a user does not have any reviews in the training split for city c, we set $z_{\text{prox}}(\cdot, u) \equiv 0$. Thus, s_{B1} reduces to $z_{\text{pop}}(i)$ and s_{B2} reduces to $(1 - \alpha)\, z_{\text{pop}}(i) + \alpha\, z_{\text{loc}}(i, c)$. For restaurants with zero reviewers in the training split, we assign $z_{\text{loc}}(i, c) = 0$, reducing s_{B2} to s_{B1}. This handling ensures a well-defined ranking under sparsity.

4. Experiments

4.1. Dataset

We utilize the Yelp Open Dataset [31], a snapshot of specific metropolitan areas that links users to businesses through ratings, textual reviews, and rich venue metadata. Following standard practice in POI recommendation, we treat each city as an independent dataset and focus on Philadelphia (PA), which has the highest review volume among available cities (967,454 reviews, compared to 635,319 for the second-ranked city, New Orleans), and restrict the domain to restaurants, as this domain provides sufficient samples for robust statistical computations.

After preprocessing operations (deduplication, missing value removal, table joins, city and domain filtering), we perform a chronological 80/20 split: the earliest 80% of interactions form the training split while the latest 20% of interactions form the test split. Table 1 summarizes the resulting entity counts.

Table 1
Dataset sizes at each preprocessing step

Stage	#Reviews	#Users	#Businesses
Raw Yelp snapshot	6,990,280	1,987,897	150,346
Philadelphia (all categories)	967,454	279,828	14,558
Philadelphia (Restaurants)	687,289	209,513	5,852
Train (earliest 80%)	549,831	164,060	5,252
Test (latest 20%)	137,458	62,244	4,082

4.2. Configuration Grid

The rankers defined in Section 3 are fixed convex combinations of standardized signals, controlled by global weights rather than learned parameters. We sweep (i) $\beta \in \{0.25, 0.5, 0.75\}$ for the tradeoff between popularity and proximity in the baseline and locality-integrated rankers, and (ii) $\alpha \in \{0.1, 0.2, \ldots, 0.9, 1.0\}$ for the locality weight in the locality-integrated ranker. We omit β=0.75 from the reported results, as the high popularity weight reduces the proximity signal to a near tiebreaker, yielding patterns qualitatively indistinguishable from β=0.5. Similarly, we report α only up to 0.4, as higher values do not yield additional insight.

4.3. Evaluation Metrics

We evaluate ranked lists at cutoffs $k \in \{5, 10, 20, 50\}$ using one accuracy and three sustainability-oriented metrics. Additionally, we define two more metrics to measure the locality orientation of recommendation lists.

- **Normalized Discounted Cumulative Gain (nDCG@k)** [32] measures ranking quality with position-aware gains, assigning higher credit to relevant items appearing at earlier positions. With binary relevance ($\mathrm{rel}(i, u) = 1$ if item i appears in user u's test interactions), we compute $\mathrm{DCG@}k(u) = \sum_{r=1}^{k} \frac{\mathrm{rel}(i_r,u)}{\log_2(1+r)}$, normalized by the ideal DCG. We report nDCG as a relevance guard to verify that locality gains do not come at a disproportionate accuracy cost.
- **Expected Popularity Complement (EPC@k)** [14] measures novelty by rewarding recommendations of less popular venues. Let $\mathrm{poprank}(i) \in \{1, \ldots, N\}$ be item i's popularity rank in the training split sorted ascending (rank 1 = least popular, rank N = most popular), and N the catalog size. Then:
$$\mathrm{EPC@}k(u) = \frac{1}{k} \sum_{r=1}^{k} \left(1 - \frac{\mathrm{poprank}(i_r)}{N}\right).$$
Higher values indicate less popular, more novel recommendations.
- **Geographic Spread (GeoSpread@k)** measures within-list spatial diversity. We define it as the mean pairwise Haversine distance (in kilometer) among all restaurants in the top-k list, capturing how geographically dispersed the recommended venues are.
- **Item Coverage (ItemCov@k)** [25] measures catalog breadth as the fraction of distinct venues that appear at least once across all test users' top-k lists. Higher values indicate that the ranker exposes a broader portion of the catalog.
- **Weighted Mean Locality (WMLoc@k)** — we define this metric to directly measure the locality orientation of recommendation lists. It computes a DCG-style rank-weighted mean of z_{loc} over the top-k list, with position weights $w_r = \frac{1/\log_2(r+1)}{\sum_{j=1}^{k} 1/\log_2(j+1)}$, formally $\mathrm{WMLoc@}k(u) = \sum_{r=1}^{k} w_r \, z_{\mathrm{loc}}(i_r, c)$. Positive values indicate the average recommended item is more locally oriented than the city mean; negative values indicate tourist-oriented bias.
- **Locality Share (LocShare@k)** — we define this as the fraction of top-k items with $z_{\mathrm{loc}}(i, c) > 0$ (i.e., more locally oriented than the city average): $\mathrm{LocShare@}k(u) = \frac{1}{k} \sum_{r=1}^{k} \mathbb{I}[z_{\mathrm{loc}}(i_r, c) > 0]$.

Given the sparse interaction setting, absolute metric values are expected to be small and should be interpreted comparatively (i.e., via relative changes across rankers) rather than in isolation.

4.4. Robustness checks

To assess sensitivity to key design choices in locality estimation, we run two analyses while keeping the remaining hyperparameters fixed. First, we vary the Empirical-Bayes prior strength ($s \in \{1, 5, 10\}$) to observe the effect of weaker versus stronger shrinkage toward the city prior, particularly for venues with few unique reviewers. Second, we vary the minimum user history threshold ($min\_user\_reviews \in \{5, 10, 20\}$) to test stability when including shorter histories or restricting to more reliable users. Based on the observations, we use $s{=}5$ and $min\_user\_reviews{=}10$ as the default setting to balance estimate stability and sample size.

5. Results & Discussion

This section reports results under the experimental setup defined in Section 4. While we compute the full grid over (α, β) and all cutoffs $k \in \{5, 10, 20, 50\}$, we report only $k \in \{10, 50\}$ and a compact subset of configurations to avoid overloading the main narrative. We focus on two popularity–proximity settings, $\beta \in \{0.25, 0.5\}$, and show the baseline $s_{\mathrm{B1}}(\beta)$ together with several locality-integrated operating points $s_{\mathrm{B2}}(\beta, \alpha)$ at $\alpha \in \{0.1, 0.2, 0.3, 0.4\}$.

Correlation check. We find that the locality score is largely independent of popularity (Pearson $r{=}0.027$, Spearman $\rho{=}0.075$), suggesting that locality contributes information orthogonal to $\mathrm{pop}(i)$ rather than acting as a popularity proxy.

Spatial distribution of locality. Figure 1 visualizes the spatial distribution of locally oriented and tourist-oriented restaurants. Dense red clusters indicate strongly locally oriented venues, while dense blue clusters indicate strongly touristic ones. Both orientations coexist in several neighborhoods, though with different density patterns, suggesting that locality-aware ranking can redistribute recommendations within these neighborhoods. The effects of this redistribution on recommendation lists are examined below.

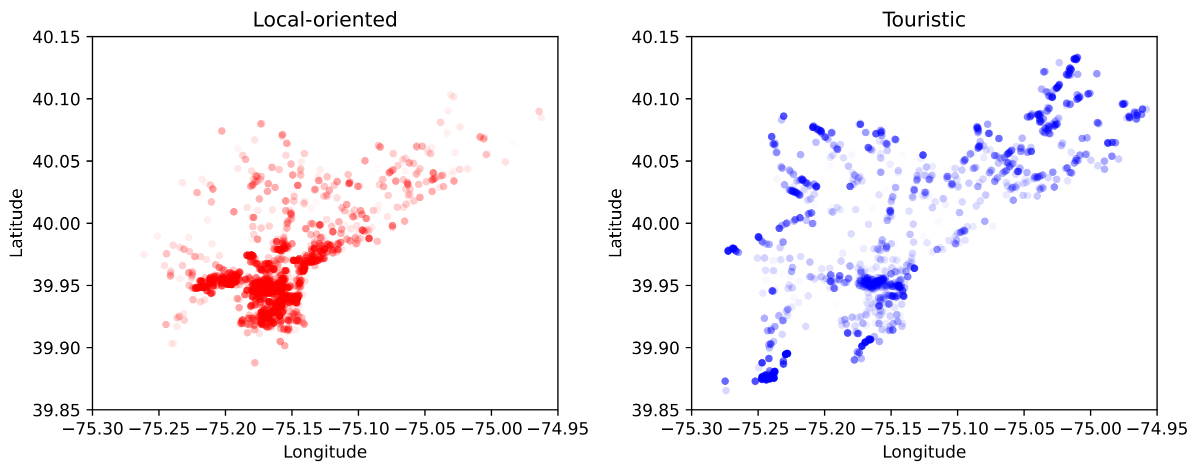

Figure 1: Spatial distribution of locally oriented (left, red) and tourist-oriented (right, blue) restaurants in Philadelphia. Color intensity reflects the magnitude of z_{loc}, with darker points indicating stronger orientation.

Results. Table 2 demonstrates that increasing the locality weight creates a clear tradeoff between nDCG and EPC in both settings, confirming that the locality score makes recommendations more novel with a cost at ranking relevance. For instance, at $\alpha = 0.2$ and k=10, EPC increases by $+142.9\%$ with a drop of -26.7% in nDCG for the balanced setting, and by $+196.0\%$ in EPC with -40.2% nDCG decrease for the proximity-forward setting. This novelty shift is directly reflected in the locality metrics: WMLoc rises from -0.81 to 0.31 and LocShare from 0.04 to 0.50, indicating that the majority of top-10 items become locally oriented. Additionally, careful tuning of α can broaden geographic spread and increase catalog coverage. Together, these findings support the sustainability argument: the locality score steers recommendations away from popular touristic hotspots toward locally oriented restaurants, distributing them more broadly across the catalog and geographically. At k=50, the trend persists and further strengthens the above findings. We also observe a clear saturation effect with increasing locality weight: while EPC continues to improve, ItemCov collapses and GeoSpread grows sharply, indicating that aggressive locality concentrates recommendations on a narrow set of highly local venues that are spatially dispersed. This reinforces that careful tuning of α is important; we find $\alpha = 0.2$ to be a good choice, as it captures meaningful sustainability gains without overspecializing the recommendation lists. Furthermore, both baselines yield negative WMLoc values across both cutoffs, indicating that popularity–proximity ranking favors tourist-oriented restaurants, further motivating the integration of the locality score.

Table 2
Results at k=10 and k=50. Bold denotes the best value per metric within each setting.

Ranker	nDCG	EPC	ItemCov	GeoSpread	WMLoc	LocShare
k=10						
Balanced setting (β=0.5)						
$B1(\beta{=}0.5)$	0.0352	0.0035	0.0194	1.43	-0.81	0.04
$B2(\beta{=}0.5, \alpha{=}0.1)$	**0.0358**	0.0038	0.0213	1.26	-0.53	0.13
$B2(\beta{=}0.5, \alpha{=}0.2)$	0.0258	0.0085	**0.0215**	1.33	0.31	0.50
$B2(\beta{=}0.5, \alpha{=}0.3)$	0.0112	0.0242	0.0124	1.91	1.29	0.83
$B2(\beta{=}0.5, \alpha{=}0.4)$	0.0026	**0.0411**	0.0055	2.31	1.96	0.99
Proximity-forward setting (β=0.25)						
$B1(\beta{=}0.25)$	**0.0328**	0.0125	0.0721	1.08	-0.71	0.12
$B2(\beta{=}0.25, \alpha{=}0.1)$	0.0291	0.0180	**0.0761**	0.94	-0.02	0.39
$B2(\beta{=}0.25, \alpha{=}0.2)$	0.0196	0.0370	0.0627	1.05	0.90	0.71
$B2(\beta{=}0.25, \alpha{=}0.3)$	0.0110	0.0558	0.0458	1.56	1.47	0.83
$B2(\beta{=}0.25, \alpha{=}0.4)$	0.0031	**0.0741**	0.0302	2.29	2.01	0.99
k=50						
Balanced setting (β=0.5)						
$B1(\beta{=}0.5)$	**0.0563**	0.0074	0.0556	1.88	-0.62	0.18
$B2(\beta{=}0.5, \alpha{=}0.1)$	0.0523	0.0085	**0.0595**	1.72	-0.26	0.36
$B2(\beta{=}0.5, \alpha{=}0.2)$	0.0411	0.0139	0.0569	1.89	0.37	0.60
$B2(\beta{=}0.5, \alpha{=}0.3)$	0.0215	0.0277	0.0409	2.29	1.15	0.85
$B2(\beta{=}0.5, \alpha{=}0.4)$	0.0083	**0.0453**	0.0201	2.44	1.71	0.97
Proximity-forward setting (β=0.25)						
$B1(\beta{=}0.25)$	**0.0534**	0.0209	0.1891	1.52	-0.49	0.24
$B2(\beta{=}0.25, \alpha{=}0.1)$	0.0453	0.0276	**0.2242**	1.35	0.09	0.51
$B2(\beta{=}0.25, \alpha{=}0.2)$	0.0313	0.0432	0.2234	1.58	0.82	0.75
$B2(\beta{=}0.25, \alpha{=}0.3)$	0.0201	0.0593	0.1947	1.97	1.31	0.85
$B2(\beta{=}0.25, \alpha{=}0.4)$	0.0090	**0.0745**	0.1233	2.38	1.74	0.96

Robustness checks. We test robustness to (i) prior strength $s \in \{1, 5, 10\}$ and (ii) minimum user history threshold $min\_user\_reviews \in \{5, 10, 20\}$. Stronger priors dampen the locality signal via heavier shrinkage, reducing EPC gains while preserving the qualitative tradeoff. Raising the history threshold to 20 reduces both accuracy and sustainability metrics due to decreased evaluation coverage. Therefore, we use $s{=}5$ and $min\_user\_reviews{=}10$ in the primary setup as a stable operating point.

Implications. A continuous, city-relative locality score provides an interpretable lever analogous to popularity and proximity. For platforms, locality can be exposed as an explicit control (or policy parameter) to adapt recommendations to user intent (e.g., familiarity-seeking versus exploration) and to mitigate excessive concentration on touristic hotspots by increasing the share of locally oriented venues in ranked lists. Beyond platform design, aggregated locality-aware recommendation outputs can support stakeholder analyses for urban planning, such as diagnosing concentration dynamics and monitoring potential "tourist bubble" effects over time. These applications require externally defined objectives and stakeholder input, but the proposed feature offers a transparent mechanism to operationalize such objectives.

6. Conclusion, Limitations & Future Work

This paper introduces a continuous, city-relative locality score, derived from review-based behavioral signals, that places venues on a spectrum from tourist-oriented to locally oriented. We integrate this interpretable item-side feature into a transparent popularity–proximity ranker via tunable weights, enabling controllable operating points between ranking relevance and sustainability-oriented objectives. In offline chronological top-k evaluation on Yelp restaurants in Philadelphia, locality-aware variants consistently increase novelty and shift recommendations toward more locally oriented venues. Robustness checks further indicate that the qualitative direction of this tradeoff remains stable under reasonable variations in key design choices (e.g., shrinkage strength and eligibility thresholds).

We note several limitations that contextualize these findings. This work evaluates on a single platform (Yelp), in a single city (Philadelphia), and within a single domain (restaurants), which limits its generalizability. Moreover, we study a classical convex-combination ranker to preserve interpretability; the magnitude and interactions of locality effects may differ under higher-capacity or learned geo-aware recommenders. Finally, our results are based on offline evaluation; real-world impacts on user satisfaction and business outcomes require validation through user studies and provider-side analyses.

Future work includes extending the scope beyond a single city, platform, and venue type by replicating the analysis across multiple cities and datasets (e.g., NYC/Tokyo check-ins [33]), and by evaluating additional POI categories. A second direction is integrating locality into stronger geo-aware recommenders (e.g., learning-to-rank or representation-based POI models) to test whether the observed sustainability gains persist under higher model capacity and to learn context-dependent operating points (e.g., by district, season, or trip intent) rather than a single global weight. Another direction is that offline metrics cannot fully capture perceived usefulness, trust, and behavioral responses; user studies and provider-side analyses, as in [34], are needed to assess the desirability of locality-aware controls and to evaluate downstream effects on user satisfaction and neighborhood-level impacts. Together, these directions would help establish locality-aware recommendation as a practical and transparent tool for promoting sustainable urban tourism, reducing concentration in touristic hotspots and broadening visitor flows across the city.

Declaration on Generative AI

During the preparation of this work, we used ChatGPT for paraphrasing and limited information search, and Grammarly for proofreading and grammar checking. After using these services, we reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] M. M. Hasnat, S. Hasan, Identifying tourists and analyzing spatial patterns of their destinations from location-based social media data, Transportation Research Part C: Emerging Technologies 96 (2018) 38–54. doi:10.1016/j.trc.2018.09.006.

[2] F. Girardin, F. Calabrese, F. D. Fiore, C. Ratti, J. Blat, Digital footprinting: Uncovering tourists with user-generated content, IEEE Pervasive Computing 7 (2008) 36–43. doi:10.1109/MPRV.2008.71.

[3] L. W. Dietz, R. Roy, W. Wörndl, Characterisation of traveller types using check-in data from location-based social networks, in: J. Pesonen, J. Neidhardt (Eds.), Information and Communication Technologies in Tourism 2019, Springer International Publishing, Cham, 2019, pp. 15–26. doi:10.1007/978-3-030-05940-8_2.

[4] G. J. Ashworth, S. J. Page, Urban tourism research: Recent progress and current paradoxes, Tourism Management 32 (2011) 1–15. doi:10.1016/j.tourman.2010.02.002.

[5] D. Li, X. Zhou, M. Wang, Analyzing and visualizing the spatial interactions between tourists and locals: A flickr study in ten US cities, Cities 74 (2018) 249–258. doi:10.1016/j.cities.2017.12.012.

[6] D. R. Judd, Constructing the tourist bubble, in: D. R. Judd, S. S. Fainstein (Eds.), The Tourist City, Yale University Press, New Haven, CT, 1999, pp. 35–53.

[7] C. Colomb, J. Novy (Eds.), Protest and Resistance in the Tourist City, Routledge, 2016. doi:10.4324/9781315719306.

[8] C. Milano, M. Novelli, J. M. Cheer, Overtourism and tourismphobia: A journey through four decades of tourism development, planning and local concerns, Tourism Planning & Development 16 (2019) 353–357. doi:10.1080/21568316.2019.1599604.

[9] A. Derdouri, T. Osaragi, A machine learning-based approach for classifying tourists and locals using geotagged photos: the case of tokyo, Information Technology & Tourism 23 (2021) 575–609. doi:10.1007/s40558-021-00208-3.

[10] C. Pasquinelli, M. Trunfio, Reframing urban overtourism through the smart-city lens, Cities 102 (2020) 102729. doi:10.1016/j.cities.2020.102729.

[11] L. Esmaeili, S. Mardani, S. A. Hashemi Golpayegani, Z. Zanganeh Madar, A novel tourism recommender system in the context of social commerce, Expert Systems with Applications 149 (2020) 113301. doi:10.1016/j.eswa.2020.113301.

[12] H. Abdollahpouri, R. Burke, B. Mobasher, Controlling popularity bias in learning-to-rank recommendation, in: Proceedings of the Eleventh ACM Conference on Recommender Systems, Association for Computing Machinery, New York, NY, USA, 2017, pp. 42–46. doi:10.1145/3109859.3109912.

[13] G. Richards, Tourism trends: The convergence of culture and tourism, Technical Report, NHTV University of Applied Sciences, 2014.

[14] S. Vargas, P. Castells, Rank and relevance in novelty and diversity metrics for recommender systems, in: Proceedings of the 5th ACM Conference on Recommender Systems (RecSys '11), ACM, 2011, pp. 109–116. doi:10.1145/2043932.2043955.

[15] J. Borràs, A. Moreno, A. Valls, Intelligent tourism recommender systems: A survey, Expert Systems with Applications 41 (2014) 7370–7389. doi:10.1016/j.eswa.2014.06.007.

[16] M. Quadrana, P. Cremonesi, D. Jannach, Sequence-aware recommender systems, ACM Computing Surveys 51 (2018) 66:1–66:36. doi:10.1145/3190616.

[17] G. Adomavicius, A. Tuzhilin, Context-aware recommender systems, in: F. Ricci, L. Rokach, B. Shapira, P. B. Kantor (Eds.), Recommender Systems Handbook, Springer, Boston, MA, 2011, pp. 217–253. doi:10.1007/978-0-387-85820-3_7.

[18] M. Ye, P. Yin, W.-C. Lee, D.-L. Lee, Exploiting geographical influence for collaborative point-of-interest recommendation, in: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, 2011, pp. 325–334. doi:10.1145/2009916.2009962.

[19] P. Sánchez, A. Bellogín, Point-of-interest recommender systems based on location-based social

networks: A survey from an experimental perspective, ACM Computing Surveys 54 (2022) 223:1–223:37. doi:10.1145/3510409.
[20] E. Cho, S. A. Myers, J. Leskovec, Friendship and mobility: User movement in location-based social networks, in: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '11), ACM, 2011, pp. 1082–1090. doi:10.1145/2020408.2020579.
[21] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, C. Mascolo, A tale of many cities: Universal patterns in human urban mobility, PLOS ONE 7 (2012). doi:10.1371/journal.pone.0037027.
[22] T. Louail, M. Lenormand, O. G. Cantu Ros, M. Picornell, R. Herranz, E. Frias-Martinez, J. J. Ramasco, M. Barthélemy, From mobile phone data to the spatial structure of cities, Scientific Reports 4 (2014) 5276. doi:10.1038/srep05276.
[23] Y. Liu, W. Wei, A. Sun, C. Miao, Exploiting geographical neighborhood characteristics for location recommendation, in: Proceedings of the 23rd ACM International Conference on Information and Knowledge Management (CIKM '14), ACM, 2014, pp. 739–748. doi:10.1145/2661829.2662002.
[24] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, S. Krishnaswamy, Rank-geofm: A ranking based geographical factorization method for point of interest recommendation, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, 2015, pp. 433–442. doi:10.1145/2766462.2767722.
[25] G. K. Patro, A. Chakraborty, A. Banerjee, N. Ganguly, Towards safety and sustainability: Designing local recommendations for post-pandemic world, in: Proceedings of the 14th ACM Conference on Recommender Systems (RecSys '20), ACM, 2020, pp. 358–367. doi:10.1145/3383313.3412251.
[26] A. Banerjee, T. Mahmudov, E. Adler, F. N. Aisyah, W. Wörndl, Modeling sustainable city trips: integrating CO_2e emissions, popularity, and seasonality into tourism recommender systems, Information Technology & Tourism 27 (2025) 189–226. doi:10.1007/s40558-024-00303-1.
[27] Y.-L. Zhao, L. Nie, X. Wang, T.-S. Chua, Personalized recommendations of locally interesting venues to tourists via cross-region community matching, ACM Transactions on Intelligent Systems and Technology 5 (2014) 50:1–50:26. doi:10.1145/2532439.
[28] R. W. Sinnott, Virtues of the haversine, Sky and Telescope 68 (1984) 159.
[29] B. Efron, C. Morris, Data analysis using Stein's estimator and its generalizations, Journal of the American Statistical Association 70 (1975) 311–319. doi:10.1080/01621459.1975.10479864.
[30] J. Lee, K. Suh, A new index to assess vulnerability to regional shrinkage (hollowing out) due to the changing age structure and population density, Sustainability 13 (2021) 9382. doi:10.3390/su13169382.
[31] Yelp Inc., Yelp open dataset, https://business.yelp.com/data/resources/open-dataset/, 2022.
[32] K. Järvelin, J. Kekäläinen, Cumulated gain-based evaluation of IR techniques, ACM Transactions on Information Systems 20 (2002) 422–446. doi:10.1145/582415.582418.
[33] D. Yang, D. Zhang, V. W. Zheng, Z. Yu, Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns, IEEE Transactions on Systems, Man, and Cybernetics: Systems 45 (2015) 129–142. doi:10.1109/TSMC.2014.2327053.
[34] P. Banik, A. Banerjee, W. Wörndl, Understanding user perspectives on sustainability and fairness in tourism recommender systems, in: Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23 Adjunct), ACM, 2023, pp. 241–248. doi:10.1145/3563359.3597442.