Exploratory Analysis of Recommending Urban Parks for Various Leisure Activities

Anonymous Author(s)*

ABSTRACT

Recommender systems could help individuals explore a city's parks. This is especially important in large cities like London having over 1,500 named parks, where understanding what different parks offer can be overwhelming. Moreover, parks are key spaces for fostering urban health. Given the lack of datasets suitable for exploring this task, and that each park could be recommended for different health-promoting activities, it is unclear which recommendation algorithms are well-suited. To investigate the dynamics of recommending parks for specific activities, we produced two datasets first, one through a survey of over 250 London residents, and another one by inferring visits from over 1 million geotagged Flickr images in London parks. Analyzing the geographic patterns of these visits, we found that recommending nearby parks is not effective, making the recommendation task distinct from the task of POI recommendation. We then tested different recommendation strategies, uncovering a high popularity bias in the results. By then removing just a small portion of the most popular parks, we found that neighborhood-based models perform the best. The data and the findings in this study serve as foundation for future work on the task of park recommendation.

ACM Reference Format:
Anonymous Author(s). 2024. Exploratory Analysis of Recommending Urban Parks for Various Leisure Activities. In *Proceedings of ACM Conference on Recommender Systems (RecSys'24).* ACM, New York, NY, USA, 10 pages.

1 INTRODUCTION

Cities are currently being transformed by urbanization [29]. As their populations grow, so does the need for spaces where people can relax and play. Parks are ideal urban spaces for such activities, and they are known to support public health and well-being [66, 68], particularly for individuals from socioeconomically disadvantaged backgrounds [37] and elderly [15]. Given the numerous advantages of parks, it is important to make them accessible to all citizens.

Understanding what each park offers can be overwhelming, especially in large cities such as London. A way of helping citizens, especially newcomers, to navigate this challenge is through recommender systems that can help to understand which parks are available for their needs. However, it is unclear what kind of recommendation algorithms are well-suited for park recommendations, so that each park may be recommended for various activities.

To tackle this gap, we introduce and study the novel task of park recommendations, with the main purpose to support urban health promotion. We treat parks as items, and state the problem as follows a user wants to do a certain activity in a park, but it is unknown which park would be suitable for this activity. The recommender system should infer the user's preferences from previous visits and recommend parks that are most suitable. Given the absence of recommendation datasets for our task, we first identified five primary categories of park activities with the potential to benefit health using a scoping review. Those activities are nature-appreciation, physical, environmental, social, and cultural [32, 37].

Then, we examined the efforts to characterize items for recommendation systems generally, and for parks more specifically. Katsumi et al. unveiled new tourist attractions by identifying lesser-known POIs with characteristics similar to those of renowned attractions [25]. Dietz et al. conducted evaluations of destinations, ensuring that the attributes' values are in line with standards set by experts [14], and subsequently integrated this framework into a conversational agent [12]. Lim et al. proposed personalized travel itineraries that utilized the popularity of POIs and the individual's travel preferences, inferred from travel patterns evident in Flickr images [35]. Zhao et al. mined interest-based communities in specific locales using Foursquare data, matching these with the interests of users [73]. Le Falher et al. characterized neighborhoods in different cities in order to recommend a neighborhood of similar characteristics to the neighborhood the user is familiar with but in an unknown city [33]. The variety of these approaches shows that characterizing items can be done in myriad ways to improve recommender systems. When it comes to parks, specifically, we did not identify research to characterize them beyond POIs, or to recommend how to explore them [51].

To go beyond previous work, we chose to characterize parks by the health-promoting activities they offer to citizens. Given that recommendation systems can nudge people to do healthy activities, we use this characterization to recommend parks for each of the five identified categories of health-promoting activities. Taking Greater London as a case study, we used OpenStreetMap (OSM) to characterize its parks for each of the five activity categories. Afterward, we created two datasets of user preferences for these activities: one using a survey of over 200 London residents, which directly asked people about suitable parks for certain activities, and another one, using 1,065,197 geotagged images in 1,463 parks in London, from which we inferred the parks users visited and for which activities.

We found that the geographic patterns of park visits differ from those of POIs more generally in a way that park visits are not highly local. This limits the adoption of traditional POI recommendation algorithms. In response to this finding, we designed a series of offline recommendation experiments based on the Elliot Framework [3] to understand the factors that influence the quality of park recommendations, especially regarding recommending different activities. In a nutshell, we made two main contributions:

(1) Benchmark datasets for recommending activities in parks (§3). To address this, we created and openly released two datasets for training and evaluating recommender systems for this task[1]. To

RecSys'24, October 2024, Bari, Italy
2024.

[1]https://anonymous.4open.science/r/park-recommendation-datasets-1975

do so, first, we profiled London parks in terms of five categories of health-promoting activities (§3.1) allowing us to use this characterization as park activity labels. We then gathered two distinct sets of user preference data (§3.2), one from a survey and the other from geo-located images on Flickr.

(2) Exploratory analysis and evaluation of existing methods (§4). We found that recommending nearby parks as done by techniques for recommending POIs is not effective. Our follow-up analysis of five widely used recommendation models revealed a significant popularity bias in the recommendations. We studied the impact of reducing this bias by excluding a small fraction of the most popular parks from the data and found that personalized recommendation models consistently outperform non-personalized ones.

2 BACKGROUND ON HEALTHY ACTIVITIES IN PARKS

Characterizing Parks. In literature outside of recommender systems, parks have been characterized in terms of when [20, 28, 36], how [20, 62], and which people [28, 36] use those parks. Chuang et al. illustrated the influence of park facilities on the composition and number of park visitors [9]. Parks have been also linked with various health benefits [5, 32, 37, 40, 72]. The mechanisms behind these benefits include physical activities [10, 23, 47], but also the mere presence in nature [44, 64], as well as socializing [11, 61], gardening [19, 42], and even cultural activities [57] in parks. Brown et al. analyzed relationships between park activities and park type, size, and location offering a deep dive into physical activity benefits of various parks [8]. However, there are no recommendation methods, which take into account varied characteristics of parks or the activities and health benefits they provide.

While some POI recommendation approaches use context information beyond location for recommendations, these approaches do not consider different functions that parks can fulfill. To characterize parks in terms of the health-promoting activities they offer, we surveyed health benefits of activities in parks.

Scoping Literature Review. To map the literature on health-promoting activities in parks, we performed a scoping review, following the widely recognized PRISMA method [46] guidelines. A scoping review was a suitable option for our task since it allows for a broad examination of the literature irrespective of the design and strength of effects found in the included studies. Parks have been suggested to promote health through various activities, such as *physical* (e.g., running), *nature-appreciation* (e.g., watching birds), *environmental* (e.g., gardening), *social* (e.g., spending time with other parents in a children playground), and *cultural* (e.g., visiting monuments and heritage). To ensure a full characterization of health-promoting activities people can do in parks, in the scoping review, we performed a set of queries on PubMed and SpringerLink databases for each of those activities using queries of the type: `(urban greenery) AND (health) AND (<exemplary activities from the category>)` in the article title or anywhere in the body. For instance, for *environmental activities* we searched for: `(urban greenery) AND (health) AND (garden OR planting OR conservation)`. An article was deemed relevant if the results evidenced that one or more activities typically done in public urban green spaces had a health benefit.

Following the PRISMA guidelines, from an initial 762 articles, we identified 344 unique articles after removing duplicates. Screening excluded 22 articles (conference proceedings and perspectives), and upon assessing eligibility, 114 articles were deemed relevant. Most articles were excluded because they were not about urban green spaces or because there was no statistically significant link between the activities and health benefits. From the included articles, we documented specific activities in each activity category, as well as specific health benefits, as illustrated in Table 1.

3 METHODS

3.1 Characterizing Parks in Terms of Activities

Characterizing Parks using OSM Tags. OpenStreetMap (OSM) is a global, crowdsourced geographic database crucial for mapping and scientific research, including, for example, urban planning [39] and disaster management [21]. It offers an extensive API used by various applications, and its open license has promoted a broad contributor base, achieving high-quality global coverage. Our study used OSM data, focusing on two types of park features (a subset of the complex OSM data model): elements (like benches), and spaces (like lakes). Each of the park elements and spaces can be tagged by contributors using key=value pairs such that the key defines a feature's category or type, and the value offers detailed information about the specified key (e.g., `natural=tree; fitness_station=horizontal_bar; artwork_type=sculpture, or garden:type=community`).

Associating Parks with Health-promoting Activities. The OSM tagging system is complex and exhibits diverse tags. We streamlined the annotation of OSM tags for park elements and spaces with health activities by focusing on the most frequent tags, and by removing a subset of tags deemed irrelevant for assessing health-promoting aspects (e.g., we removed tags related to the `material` and `inscription` of a bench). Due to the impracticality of expert or crowdsourced annotation for over 30,000 unique OSM tags found in parks, we utilized and validated LLM classifiers. We did so in four steps. First, we built the ground-truth list from the 100 most frequent tags, which, because of their frequency, covered a large portion of all park-tag pairs. These tags were manually annotated by our team. To ensure accuracy and reliability, three co-authors independently labeled these 100 items with health-promoting activities or *"none,"* and we used the majority voting strategy to aggregate the individual opinions into one final outcome label. In cases where conflicts arose, i.e., where the three annotators provided different labels, a discussion was held to resolve the discrepancies. In the resulting ground-truth dataset, for instance, the tag `natural=tree` was associated with *nature-appreciations*, `garden:type=community` with *enviromental*, and `artwork_type=sculpture` with *cultural* activities. This list was used for the LLM benchmarking. Second, we crafted a prompt asking to label each OSM tag with health-promoting activity category, providing a brief definition of the OSM tags taken from the OSM wiki, and assigning a role of *"an expert in urban planning and public health, with a specialization in urban parks,"* and compared two LLM models (`gpt-3.5-turbo` and `gpt-4`) with temperatures $t \in \{0.3, 0.6, 0.9\}$, controlling the randomness of the models' completions. We found that GPT-4 set at a temperature of 0.9 achieved the best performance on the benchmark. Third, we

Table 1: Example activities and specific health benefits discovered for each activity category during our literature review. The last column shows exemplary OSM tags we associated with each of the activity categories.

Activity Cat.	Example Activities	Health Benefit	OSM Tags
Physical	hiking, trail running, biking, swimming, rock climbing, canoeing	weight reduction [34], cardiovascular health improvements [49], social cohesion [7]	`leisure=pitch, leisure=playground, sport=fitness, sport=soccer`
Nature-appreciation	bird watching, picnicking, fishing, painting	positive emotions [43], mood improvement [71], increased social capital [18]	`natural=tree, amenity=fountain, natural=wood, natural=water`
Environmental	gardening, planting trees and flowers, and participating in conservation efforts	stress reduction [19], access to healthy produce [74], social cohesion [60]	`waste=trash, produce=plum, leisure=garden, landuse=flowerbed`
Social	playing sports, kids playgroups, volunteering	mood improvement [31], dementia prevention [57], sense of social belonging [56]	`amenity=bench, tourism=information, leisure=picnic_table, amenity=cafe`
Cultural	festivals, art exhibits, music performances	dementia prevention [57], increase of physical activity [48], quality of life [57]	`tourism=attraction, tourism=artwork, historic=memorial, leisure=bandstand`

evaluated the result of the best-performing model by asking three domain experts independently (and aggregating their responses using majority voting) whether the top 20 most frequent annotations for each activity category were correct, finding the overall accuracy of 92%. Lastly, employing the literature-derived taxonomy of five health-promoting activity categories (Table 1) and GPT-4 tag annotation, we categorized all OSM park elements and spaces according to how their tags were labeled, assigning *"none"* where no clear activity support existed. Hence, park elements and spaces were associated with health activities (e.g., forest with nature-appreciation, community gardens with environmental, and monuments with cultural activities).

3.2 Collecting Park Recommendation Datasets

Parks as items in recommender systems are not comparable to books, movies, music, nor, as we will show, POIs. The reason for this is that parks are urban spaces with a wide array of offerings to different users and even the same user in different situations, i.e., when they want to do different activities, such as sports or enjoying nature. Using the characterization of parks in terms of their offerings for health-promoting activities, we established a richer understanding of parks as recommendation items. To be able to evaluate approaches for activity-aware park recommendations, we first needed a dataset to perform recommendations.

As such data does not exist, we established two complementary data sets for our evaluations. The first is a survey-based dataset, where we asked citizens to name parks they find suitable for doing certain activities. The second is based on a large-scale image data set from Flickr (https://flickr.com), which we processed into an implicit feedback recommendation dataset of park visits.

Survey on Parks Suitable for Leisure Activities. In an online survey, we asked London citizens about suitable parks for performing each of the five types of activities from our taxonomy. The main question for physical activities was phrased as: *"Can you name several parks suitable for physical activities (e.g., sports)?"*, and we repeated a similar phrasing for other activities.

We recruited the participants using the first author's institutional research recruitment portal (n = 81), as well as through Prolific[2] (n = 178). The participants (F: 48%, M: 41%, O: 11%) were informed about the purpose of the survey, the voluntary nature of their participation, and the scope of the data collection. The average age was 36.2 ± 10.94 and the home locations of the respondents was uniform across the Boroughs of London, with the exception that only few were from the centrally located City of London. We asked participants for their home postal area (e.g., N1) and how long they have been living in London (*"I don't live in London."* – *"Less than 1 year."* – *"1 to 5 years."* – *"More than 5 years."*). Finally, as a means to identify low-quality responses, we asked people for a park close to their homes, which we could use as an instructional manipulation check in conjunction with the reported postal area. In the Prolific study, as a second attention check, in one of the steps, we asked the respondents to simply select 'Hyde Park' instead of naming parks suitable for a certain activity. The data collection was registered as a minimal-risk study at the first author's institutional review board. After having removed the answers of users who failed any attention check, we interpreted each park of the survey as a signal that the user finds it suitable for the respective activity, thus, we obtained a recommendation dataset where each mention of a park leads to an entry in the user-item matrix. To establish a uniform terminology, we use the term *visit* for each mention by a participant of a park in the rest of the paper. The characteristics of the survey are summarized in Table 2.

Flickr. It is one of the most prominent platforms for sharing photography. Since its inception in 2004, Flickr has gained considerable popularity, accumulating billions of images. Notably, many of these images have been precisely geo-located, thanks to the utilization of the (phone) camera's GPS module. We utilized a substantial dataset, collected using Flickr's API, comprising geo-located images posted between 2004 and 2015. This extensive dataset offered us a valuable resource for park visits in London. By intersecting the 12 million images with the park outlines as defined in OpenStreetMap, we identified 1,065,197 individual visits to 1,463 parks. Note that we only included named parks in the park list, as unnamed parks are quite small in size, are typically not under active maintenance and without a name are cumbersome to refer to in a recommender system. Following several pre-processing steps, we converted these visits into a useful and realistic dataset for park recommendations:

In the first step, we selected all images of a user within London and computed the centroid from these to estimate the user's center of life within the city, which has been shown to approximate the home location [50]. In the next step, we used OpenStreetMap

[2] https://www.prolific.com/

Table 2: Characteristics of the datasets (survey left, Flickr, right). VPU – visits per user, VPP – visits per park.

Survey	Users	Parks	Visits	Density (in %)	VPU	VPP
Cultural	208	97	501	2.4831	2.4087	5.1649
Env.	178	125	347	1.5596	1.9494	2.7760
Nature	249	146	630	1.7330	2.5301	4.3151
Physical	252	203	947	1.8512	3.7579	4.6650
Social	248	156	882	2.2798	3.5565	5.6538

Flickr	Users	Parks	Visits	Density (in %)	VPU	VPP
Cultural	6132	655	10426	0.2596	1.7003	15.9176
Env.	420	189	490	0.6173	1.1667	2.5926
Nature	10203	821	17692	0.2112	1.7340	21.5493
Physical	4743	628	7202	0.2418	1.5184	11.4682
Social	3919	510	5626	0.2815	1.4356	11.0314

to gather the shapes of all parks in London and discarded all images that were not within the park boundaries. This left us with 1,065,197 individual park visits. As we are interested in specific activities of the user, we analyzed two types of tags the images were partially annotated with: user-generated tags and computer vision labels from a computer vision algorithm [63]. To match these tags to activities, we employed Sentence-BERT (S-BERT [53]) for text embeddings. For this, we utilized the OSM tags of amenities and spaces found in parks, which were already annotated with an activity each (cf. § 3.1, *Associating Parks with Health-promoting Activities through Their Facilities*). After embedding the OSM tags using the `all-mpnet-base-v2` model, we matched each Flickr label to the closest OSM tag in the embedding space, using the cosine distance as similarity measure while using a matching quality threshold of a cosine similarity of at least 0.7. This allowed us to avoid matching Flickr labels that did not have meaningful OSM counterparts. Similarly to the LLM annotations (Step 3 in § 3.1), we measured the agreement of the top 20 matchings from Flickr tags to activity categories. The agreement was 82%, which is highly accurate given that the matchings are only based on individual tags.

The outcome was that we annotated all labels attached to an image with an activity. For example, if the labels were [`'sunshine'`, `'volleyball'`, `'pond'`], we would discard sunshine, as it is unrelated to any activities, and count the other to tags towards physical and nature-appreciation, respectively. The outcome of this image would be a relative frequency count of 0.5 for physical and 0.5 for nature-appreciation. Doing so for all images, we can now subdivide all images into visits to parks and infer the activity that the user was involved in. To prevent recording spurious activities, we only assign an image to an activity if the relative frequency count is at least 0.5. So, our previous example would be counted both towards physical activities and nature-appreciation activities, but if another tag were present related to physical activities, we would not count this image towards nature-appreciation anymore, as only 1 of 3 labels is matched to this activity.

In this way, we obtained an implicit feedback dataset from Flickr, which records visits of a user to parks and also contains information about which activity the user was interested in when capturing the photo. As we are interested in recommending novel items to a user, we remove all duplicate visits to a park for one activity. Table 2 shows the data characteristics of the Flickr dataset. The density of the Flickr dataset is very low, with ranges from 0.21% (nature-appreciation) to 0.62% in the environmental category, albeit with only 189 distinct park visits recorded for this activity. The city-level density is comparable with POI recommendation datasets such as Foursquare and Gowalla [59].

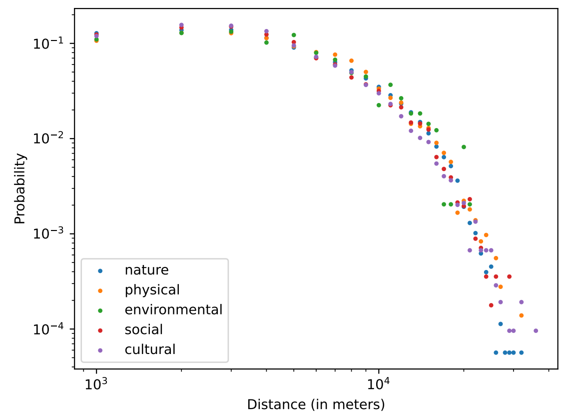


Figure 1: Probability of one's traveling a certain distance for visiting a park for each activity category. All categories follow a similar distribution with varying scaling factors (Table 3). The plot shows the distribution in the Flickr dataset.

3.3 Exploring Geographic Patterns of Visits

Analyzing the geographical aspect of visited parks reveals different patterns of visiting parks for distinct activity categories. To reveal geographic patterns in the two datasets for the individual activities, we fit a scaling factor α for the distances between the users's home/center location and the visited parks [58]:

$$p_{\text{close}} = \frac{1}{d_{u,p}^{\alpha}}, \tag{1}$$

where $d_{u,p}$ is the distance between the user's location (home postcode in the survey, center of geographic interest in Flickr) and the centroid of the visited park. To derive the probability distribution, we used a bin width of 1km. The probabilities are plotted in Figure 1.

Table 3: Statistics describing how far the visited parks were away from the user's home (Survey) or center of geographic interest (Flickr). α indicates the scaling factor of the probability of traveling a certain distance to visit a park (Equation 1). The greater the value of α, the less significance distance holds in selecting a park of that type to visit [58].

Activity Cat.	Flickr α	Survey α	Flickr Mean Distance (km)	Survey Mean Distance (km)
Cultural	1.90	2.33	4.59	7.21
Environmental	2.03	2.06	5.15	6.39
Nature	2.02	2.22	5.00	7.50
Physical	2.04	1.99	5.22	6.10
Social	1.94	2.12	4.75	6.57

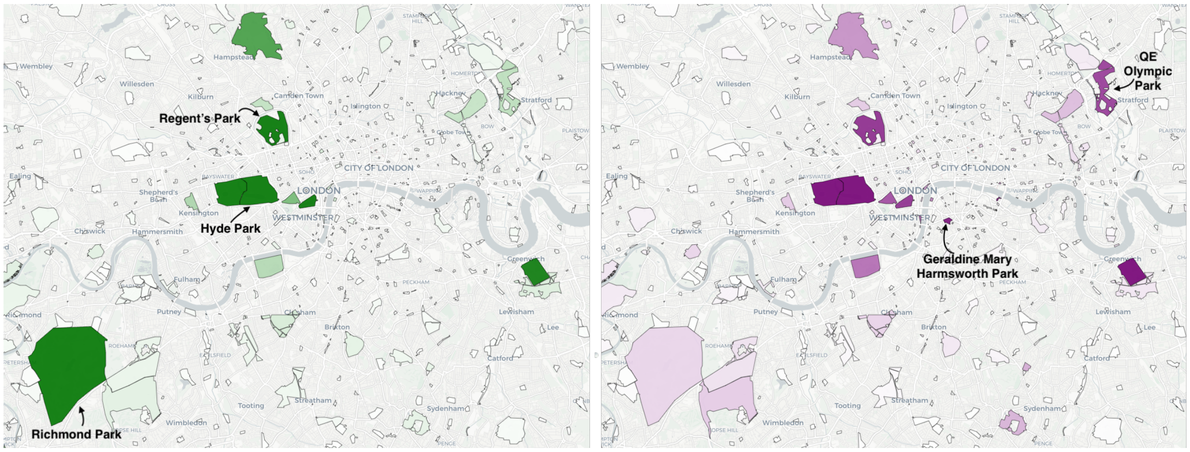


Figure 2: Maps of London showing the prevalence of nature-appreciation (left) and cultural activities (right) in the Flickr dataset. Central London parks have many visits for both activities. Richmond Park is more visited for nature-appreciation activities, whereas Geraldine Mary Harmsworth Park, which hosts the Imperial War Museum, is a distinctly cultural park.

We revealed that the distances are generally smaller in the Flickr dataset compared to the survey. By fitting the scaling factor α for the probability of traveling a certain distance for a certain activity, we found that using the center of geographic interest is likely not a good proxy for home location as the resulting values were smaller compared to the survey and inconclusive regarding the distances traveled to visit a park for a certain activity. Hence, we focused on the results in the survey dataset. The higher the value of α, the less important distance becomes in choosing to visit a park for this activity [58]. In the survey, we found that α is the highest for cultural activities and lowest for environmental and physical activities. These results make sense as there are relatively fewer parks that offer high-quality cultural experiences, such as museums, artwork, and concerts. Moreover, specific cultural activities in such parks might be organised once-in-a-while, and not continually. Hence, people are willing to travel longer distances for cultural activities taking place in a park. On the other hand, activities such as sports are part of people's busy everyday life, and are preferably done at some of the neighboring parks, and community gardening is also something people typically do within their neighborhood, as this activity requires frequent, often daily, visits.

These findings revealed the advantages and shortcomings of the two methods for establishing park recommendation datasets. With over 41,436 visits, the Flickr dataset was about 12 times larger than the survey (3,453), however, the distribution of activities is more imbalanced with very few images depicting environmental activities. The other shortcoming of the Flickr dataset was that the user location must be inferred from the center of geographic interest, which gravitates towards the city center making the geographic analysis less reliable compared to the survey.

3.4 Selecting Recommendation Models

To understand park recommendations, we needed models that are: *(i)* suitable for small datasets (since the number of parks in a city limits the dataset size), *(ii)* interpretable (enabling straightforward explanations for recommendation performance), and *(iii)* perform competitively [16]. We chose the following models:

MostPop. This non-personalized model recommends parks by popularity. The popularity is measured by the number of users that have visited the given park for the given activity. It is one of the simplest, yet effective baselines in sparse datasets [6].

User-KNN. This non-normalized KNN algorithm recommends parks that similar users visited [2].

Item-KNN. This non-normalized KNN recommends parks similar to the ones the user has visited [2]. Both, Item and User-KNN models are conceptually simple methods, yet competitive with deep learning [16].

SVD. Pure Singular Value Decomposition Recommendations [27]. The user-item matrix is projected into a latent space, which allows to estimate scores for missing user-items pairs. Thus, it is the simplest matrix factorization approach.

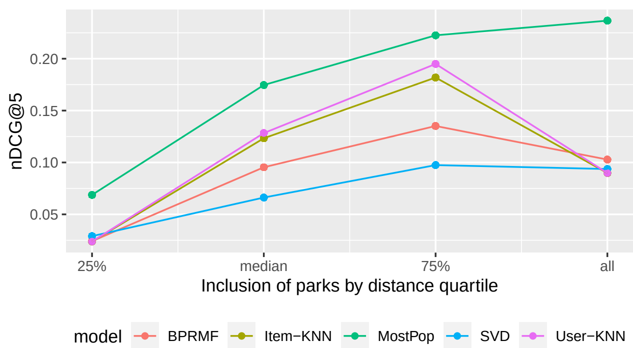
BPRMF. A matrix factorization algorithm that uses Bayesian Personalized Ranking loss as an optimization algorithm for the model parameters [54]. It was included due to its explicit fit for implicit data and its successful application in many domains [59, 70].

We also considered the latest deep learning-based recommendation algorithms and established tensor factorization techniques [22, 24]. However, these were excluded due to their black-box nature or unsuitability for smaller datasets, as the size of the datasets needs to scale with model complexity [67]. We also disregarded POI-specific algorithms, as we show that the geographic assumptions mismatch.

4 EVALUATION

The following experiments are conducted using Elliot [3]. We chose this experimentation framework as it already provided us with a large set of recommendation models, evaluation metrics, and preprocessing steps, and it offers means for reproducible evaluation.

Figure 3: Recommendation accuracy depends on how many distant parks are included in the dataset (survey only).



4.1 Goal

The overarching goal was to find out which algorithms and strategies lead to the best recommendation quality for recommending parks for a specific leisure activity. We treat this problem as an implicit feedback recommendation problem and our experiments aim to uncover what are the most influential factors to compute high-quality park recommendations.

4.2 Setup and Execution

Due to the small size of the datasets, we performed a leave-one-out split on a user basis, where one random visit per activity was reserved for the test set. A random or temporal split of a certain number of interactions per user would have been preferable. However, considering the data characteristics (Table 2), particularly the low number of visits per user, this was the only viable option to evaluate the park recommendations. Users might have visited the same park more than once, but since we wanted to recommend new parks, we discarded all duplicate check-ins of a user. Recall that the goal of a recommender system is to recommend new parks for users to explore, not parks that the user already knows.

4.3 Results

In our exploratory analysis of the two park recommendations datasets, we used a three-step approach. Each step looked at park recommendations from different angles: geographic influences, popularity bias, and the performance of recommendation models.

(1) Recommending nearby parks (as done by techniques for recommending POIs) is not effective. In a first step, we followed up on the analysis of geographic influences in §3.3 and analyzed how the recommendation accuracy is influenced by including increasingly more distant parks from the user's home. For each user, we subdivided parks into 4 quartiles based on the distance to the users' home location. To ensure using an accurate home location in this experiment, we only used the survey dataset, where this information was available. We now run the recommendation experiments four times, each time with one more quartile of parks being available for recommendation. For example, the median quartile indicated on the x-axis in Figure 3, denotes that all parks within median distances are utilized as the candidate pool for recommendations. Note that the leave-one-out split of training and test data was done after the distance-split and we additionally randomly subsampled the number of interactions to be equal to the first quartile. We repeated the experiments 100 times to average out the effects of the subsampling and leave-one-out splits.

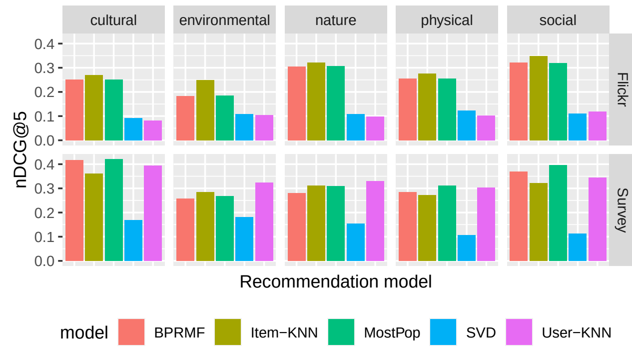


Figure 4: Recommendation accuracy (nDCG@5) in Flickr and survey datasets. We find that recommendation models influenced by popularity exhibit high performance.

Figure 3 shows that, with increasingly more distant parks in the candidate pool, we observe an increase in recommendation accuracy until the 75% quantile, and smaller gains or even worsening afterward when including more distant items. Especially, the first quartile of parks seems to include very few relevant parks that can be recommended with abysmal recommendation accuracy of an nDCG@5 < 0.05. Overall, MostPop performs best, while User-KNN slightly outperforms Item-KNN. The matrix-factorization algorithms BPRMF and SVD follow a similar pattern, but seem to suffer from the small number of interactions compared to the other models. This finding is highly relevant for the upcoming experiments, as it shows that including distant parks up to Q3 improves the accuracy of the recommendations, which violates the basic assumption of POI recommendation models that incorporate geographic proximity into their scoring [59].

(2) Recommending default (popular) parks is effective. We now turn our attention to recommending parks to users for certain activities. By subdividing the test set interactions by activity, we obtain individual results for each activity category (Figure 4). In the Flickr dataset, the item-based KNN method performs best for all activities, closely followed by MostPop and BPRMF. SVD and User-KNN perform very consistently in the range of 0.08 − 0.1 for nDCG@5. Activity-wise, the highest accuracy is achieved in nature-appreciation and social activities. In the Survey dataset, MostPop is best for cultural, physical, and social activities, but is beaten in the environmental and nature-appreciation categories by User-KNN. BRPMF and Item-KNN are also competitive leaving SVD as the only model that fails to produce high-quality recommendations. In this dataset, the highest accuracy is achieved in the recommendations for cultural activities, followed by social activities.

To assess the Popularity Bias of the recommended items, we follow the commonly accepted metric proposed by Abdollahpouri et al. [1]. We average the popularity bias over all users as per [69] resulting in the Average Ranking Popularity (ARP) metric.

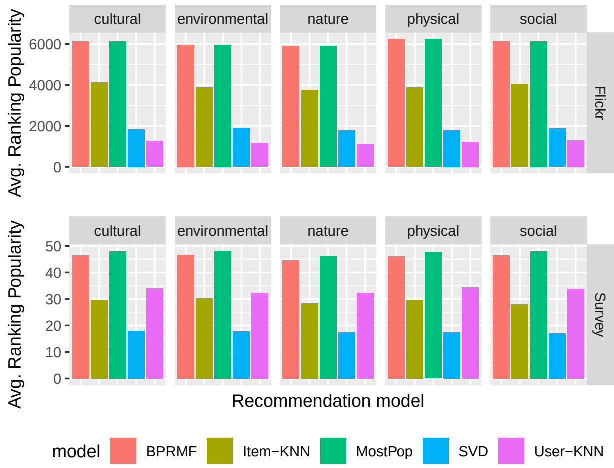


Figure 5: Average Ranking Popularity of the recommendations. We find that BPRMF and MostPop heavily rely on highly-popular parks in their recommendations raising questions about the usefulness of these recommendations.

$$\text{ARP} = \frac{1}{|U_t|} \sum_{u \in U_t} \frac{\sum_{i \in L_u} \phi(p)}{|L_u|}, \tag{2}$$

where $\phi(p)$ is the popularity scoring function for a park p. A park's popularity score is defined as the number of users who visited p over the entire number of users. U_t is the number of users in the test set and L_u is the recommended list of parks for user u. When contrasting the recommendation accuracy in Figure 4 with the popularity bias in Figure 5, we see a clear trend that MostPop and BPRMF recommend items that are highly popular, and are, thus, highly successful in both recommendation datasets. Similarly, User-KNN's success in the Survey seems to be based on recommending popular items. The exception is the Item-KNN, which offers good recommendations, especially in the Flickr dataset without relying on overly popular items.

Even though we did not allow repetitive visits to parks, the fact that the recommendations are biased towards popular items is concerning, as the value of recommending such well-known items is small, as users are likely to know them anyway. This phenomenon of popularity bias is prevalent in POI recommendation [1, 59], where a common strategy is to remove a certain portion of the most popular items to increase the value of the resulting recommendations to the user, by recommending more novel and diverse items [13, 59]. In that vein, we experiment with removing the most popular parks to mitigate popularity bias.

(3) After disregarding highly popular parks, personalized recommendation models consistently outperform non-personalized models. Removing the most popular parks, i.e., those visited by the highest number of different users, leads to a more difficult recommendation problem as the sparsity issue is further increased. However, the recommendation problems become more realistic and the usefulness of the recommendations for the user typically increases [17].

Experimenting with dropping different percentages of the most popular parks, we chose 0.5% as the threshold, which despite only removing two (Regent's Parks and Hyde Park) and eight (Hyde Park, Kensington Gardens, St. James's Park, Greenwich Park, Regent's Parks, Green Park, Jubilee Gardens, and Richmond Park) parks in the Survey and Flickr datasets, respectively, has a great effect in diminishing the popularity bias. Removing even more parks would lead to losing disproportionately many visits due to the long-tailed distribution of visits per park. The reduction in Popularity Bias is reported in Table 4. In the Flickr dataset, the reduction ranges from −42% to −84%, with Item-KNN showing the highest reduction. In the Survey dataset, the reduction is between 10% and 30%, with Item-KNN recommendation again reduced most severely by around 55%.

Revisiting the resulting accuracy metrics in the reduced dataset (Figure 6), we observe the expected reduction in accuracy but now Item-KNN and User-KNN outperform MostPop and BPRMF in Flickr and the Survey, respectively. We interpret this finding as follows: removing the *obvious* parks leads to a more realistic recommendation problem, and the user and item-based neighborhood models are best suited for these small but sparse cold-start recommendation problems [30]. The success of the Item-KNN model in the Flickr dataset can be explained by the high number of visits per park (Table 2), whereas the User-KNN works better in the Survey due to the higher number of visits per user. Surprisingly, SVD was consistently outperformed, which appears to be a consequence of the small size of the datasets coupled with their sparsity.

Table 4: Change (in %) of the Average Recommendation Popularity in the Flickr / Survey datasets when dropping the top 0.5% most popular parks. The popularity in Item-KNN decreases most when the most popular parks are dropped.

	Cultural	Env.	Nature	Physical	Social
MostPop	-67 / -20	-66 / -22	-66/ -22	-68 / -20	-66 / -21
BPRMF	-67 / -21	-66 / -23	-66 / -26	-68 / -23	-67 / -23
Item-KNN	**-84 / -58**	**-81 / -55**	**-84 / -56**	**-84 / -56**	**-84/ -56**
User-KNN	-63 / -27	-59 / -21	-61 / -28	-64 / -28	-64 / -27
SVD	-44 / -11	-45 / -12	-42 / - 9	-42 / -12	-44 / -10

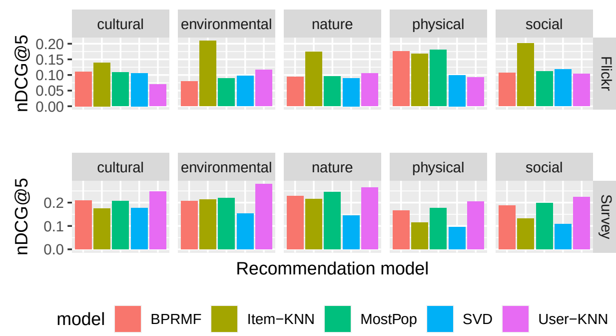


Figure 6: Recommendation accuracy of the recommendations after dropping the 0.5% most popular parks. The popularity-influenced methods (BPRMF & MostPop) are outperformed by neighborhood-based models in most activities.

5 DISCUSSION

Investigating a specific recommendation task such as recommending parks for health-promoting activities presents challenges due to limited available datasets, and uncertainty about the factors influencing successful recommendation strategies. Park recommendations serve as an important case study, particularly in large cities, as they can facilitate urban exploration and support decision-making regarding activities conducive to urban health (Table 1). From the perspective of urban planners and healthcare providers, park recommendations can be a cost-effective way of complementing initiatives towards urban health. For example, park recommendation features could be implemented in mental health apps [38] or social prescribing programmes [41]. Moreover, understanding which parks people prefer for specific health activities (e.g., physical exercise, socializing, nature appreciation) in various urban and national contexts can guide urban planners in making targeted interventions. For example, planners could promote physical activity parks in socio-economically deprived areas and parks for social and nature appreciation in areas with a higher elderly population.

To understand the dynamics of park recommendations, we focused on established and interpretable recommendation models. The User/Item-KNN proved to be strong baselines, which is in line with previous analyses, which attested these neighborhood-based models to be competitive in various domains, even in comparison with advanced deep learning recommendations [16]. BPRMF, SVD, and other matrix factorization methods likewise have been the core of successful recommendation models in new recommendation domains, e.g., in the RecSys Challenge [70].

We see the potential to adapt certain context-aware approaches to incorporate activities similar to contextual factors. Established concepts like context-aware Factorization Machines [55] or Field-aware Factorization Machines [22] might be conceptually worthwhile to consider, but their application did not yield improvements in performance, likely due to the limited size of the datasets.

Another common approach is hybrid recommendation strategies, which combine interaction data with content-based filtering using park characteristics. To gain an initial understanding of park visitations and preferences, we adopted two approaches to dataset collection. While we consider our survey dataset highly accurate, respondents may have been influenced by falling back to a default option when struggling with recall, i.e., to famous parks [52]. Furthermore, one user typically provided just around ten parks for five activities leading to a scaling issue as we are constrained by the number of people one can survey in a city. To overcome this, we created a "check-in" dataset for parks using activities from Flickr based on image content and metadata. This approach yielded a greater number of interactions, but required the inference of users' home locations, potentially introducing errors into the geographic aspects of the dataset compared to the survey method.

5.1 Limitations and Future Work

Our work comes with five main limitations. First, our novel datasets for assessing park suitability based on the survey and Flickr images do not measure actual user activities in parks. Future research could develop improved datasets, e.g., by inferring activities from mobile app activity. Analyzing the data characteristics, including the geographic patterns of such data sets will reveal further insights into the dynamics of park recommendations.

Second, the small size and specific characteristics of our datasets limit the employability of certain recommendation algorithms. This restricts the use of advanced models like deep or reinforcement learning, which need larger datasets, mirroring challenges faced in offline POI recommender systems [59]. Despite these challenges, establishing park recommender systems in practice seems feasible, as many other cities than London will have significantly fewer parks, allowing for easier data collection. Concretely, the size of datasets is limited by definition as the number of parks in a city is bounded to typically a few hundred.

Third, our reliance on offline experimentation limits insights into the user side of park recommendations. While this was insightful regarding model choices, future research should test these algorithms in real-world settings focus evaluations on practical indicators like park visit frequency and activity diversity, and explore the interdependencies of activities within parks. Finally, user-experience aspects, like presentation and reception of recommendations, will be crucial for real-world effectiveness and user adoption [26]. Designing and deploying a health-aware park recommender system in partnership with municipalities will yield valuable insights into the design space and the extent to which positive behavioural change can be realized through recommender systems.

Fourth, an ultimate aim of park recommendations is to enhance urban health, especially for vulnerable communities. The accessibility of parks and other urban green spaces for the population is part of the UN Sustainability Development goals [65] and it has been shown that single imdicators, such as the distance is not sufficient to explain the accessibility of parks [4]. Thus, it is yet to be investigated how park recommendations can be adapted for different groups, such as the elderly or economically disadvantaged to visit a suitable parks. Future work should consider socio-demographic aspects and various accessibility factors [4].

Lastly, while our study focused on one city, our methodology can be adapted to other cities, i.e., by training specific recommender models for each city. The activity taxonomy is universal, and park activity labels can be collected across the cities in which OSM is available, allowing for the adaptation of park activity labels and recommender models to local preferences and cultures [45].

6 CONCLUSIONS

We investigated the task of park recommendation, which differs from POI recommendation in two key aspects. Each park supports various health-promoting activities, which we confirm using a scoping survey. Furthermore the geographic factors affect park recommendations differently from standard POIs recommendation. After collecting and publishing two datasets, we analyzed park recommender system requirements, noting the significance of distant parks and challenges with sparsity and popularity biases of park visits. Neighborhood models performed well in recommendation experiments, with User-KNN and Item-KNN outperforming alternatives in the survey and Flickr datasets, respectively. Due to dataset size, sophisticated models like SVD underperformed, and newer techniques like deep learning or reinforcement learning were not trainable.

REFERENCES

[1] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. 2019. Managing Popularity Bias in Recommender Systems with Personalized Re-ranking. In *FLAIRS*. 413–418.

[2] Fabio Aiolli. 2013. Efficient top-n recommendation for very large scale binary rated datasets. In *RecSys*. ACM, 273–280.

[3] Vito Walter Anelli, Alejandro Bellogín, Antonio Ferrara, Daniele Malitesta, Felice Antonio Merra, Claudio Pomo, Francesco Maria Donini, and Tommaso Di Noia. 2021. Elliot: A Comprehensive and Rigorous Framework for Reproducible Recommender Systems Evaluation. In *SIGIR*. ACM, 2405–2414.

[4] Alice Battiston and Rossano Schifanella. 2024. On the need for a multi-dimensional framework to measure accessibility to urban green. *npj Urban Sustainability* 4, 1 (March 2024). https://doi.org/10.1038/s42949-024-00147-y

[5] Joshua W. R. Baur and Joanne F. Tynon. 2010. Small-Scale Urban Nature Parks: Why Should We Care? *Leisure Sci.* 32, 2 (Feb. 2010), 195–200.

[6] Alejandro Bellogín, Pablo Castells, and Iván Cantador. 2017. Statistical biases in Information Retrieval metrics for recommender systems. *Inf. Ret. J.* 20, 6 (July 2017), 606–634. https://doi.org/10.1007/s10791-017-9312-z

[7] Amy M. Bohnert, Laura M. Nicholson, Laurel Mertz, Carolyn R. Bates, and Dana E. Gerstein. 2021. Green schoolyard renovations in low-income urban neighborhoods: Benefits to students, schools, and the surrounding community. *Am. J. Commun. Psychol.* 69, 3-4 (Oct. 2021), 463–473.

[8] Greg Brown, Jonathan Rhodes, and Marie Dade. 2018. An evaluation of participatory mapping methods to assess urban park benefits. *Landscape Urban Plan.* 178 (2018), 18–31.

[9] I-Ting Chuang, Francisco Benita, and Bige Tunçer. 2022. Effects of urban park spatial characteristics on visitor density and diversity: A geolocated social media approach. *Landscape Urban Plan.* 226 (2022), 104514.

[10] Deborah A Cohen, J Scott Ashwood, Molly M Scott, Adrian Overton, Kelly R Evenson, Lisa K Staten, Dwayne Porter, Thomas L McKenzie, and Diane Catellier. 2006. Public Parks and Physical Activity Among Adolescent Girls. *Pediatrics* 118, 5 (Nov. 2006), e1381–e1389.

[11] Daniel Cox, Danielle Shanahan, Hannah Hudson, Richard Fuller, Karen Anderson, Steven Hancock, and Kevin Gaston. 2017. Doses of Nearby Nature Simultaneously Associated with Multiple Health Benefits. *Int. J. Env. Res. Pub. He.* 14, 2 (Feb. 2017), 172.

[12] Linus W. Dietz, Sameera Thimbiri Palage, and Wolfgang Wörndl. 2022. Navigation by Revealing Trade-offs for Content-based Recommendations. In *ENTER*. Springer, Cham, 149–161.

[13] Linus W. Dietz, Pablo Sánchez, and Alejandro Bellogín. 2023. Understanding the Influence of Data Characteristics on the Performance of Point-of-Interest Recommendation Algorithms. (2023). arXiv:2311.07229 [cs.IR]

[14] Linus W. Dietz, Mete Sertkan, Saadi Myftija, Sameera Thimbiri Palage, Julia Neidhardt, and Wolfgang Wörndl. 2022. A Comparative Study of Data-driven Models for Travel Destination Characterization. *Front. Big Data* 5 (April 2022).

[15] Angel Mario Dzhambov and Donka Dimitrova Dimitrova. 2014. Elderly visitors of an urban park, health anxiety and individual awareness of nature experiences. *Urban For. Urban Gree.* 13, 4 (2014), 806–813.

[16] Maurizio Ferrari Dacrema, Paolo Cremonesi, and Dietmar Jannach. 2019. Are we really making much progress? A worrying analysis of recent neural recommendation approaches. In *RecSys*. ACM, NY, USA, 101–109.

[17] Mouzhi Ge, Carla Delgado-Battenfeld, and Dietmar Jannach. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *RecSys*. ACM, NY, USA, 257–260.

[18] Peter P. Groenewegen, Jan-Paul Zock, Peter Spreeuwenberg, Marco Helbich, Gerard Hoek, Annemarie Ruijsbroek, Maciej Strak, Robert Verheij, Beate Volker, Geeke Waverijn, and Martin Dijst. 2018. Neighbourhood social and physical environment and general practitioner assessed morbidity. *Health & Place* 49 (Jan. 2018), 68–84.

[19] Jiaqi Gu, Haixiao Liu, and Hong Lu. 2022. Can Even a Small Amount of Greenery Be Helpful in Reducing Stress? A Systematic Review. *Int. J. Env. Res. Pub. He.* 19, 16 (Aug. 2022), 9778.

[20] Vuokko Heikinheimo, Henrikki Tenkanen, Claudia Bergroth, Olle Järv, Tuomo Hiippala, and Tuuli Toivonen. 2020. Understanding the use of urban green spaces from user-generated geographic information. *Landscape Urban Plan.* 201 (Sept. 2020), 103845.

[21] Benjamin Herfort, Sven Lautenbach, João Porto de Albuquerque, Jennings Anderson, and Alexander Zipf. 2021. The evolution of humanitarian mapping within the OpenStreetMap community. *Sci. Rep.* 11, 1 (Feb. 2021).

[22] Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. 2016. Field-aware Factorization Machines for CTR Prediction. In *RecSys*. ACM, NY, USA.

[23] Andrew T Kaczynski, Luke R Potwarka, and Brian E Saelens. 2008. Association of Park Size, Distance, and Features With Physical Activity in Neighborhood Parks. *Am. J. Public Health* 98, 8 (Aug. 2008), 1451–1456.

[24] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *RecSys*. ACM, NY, USA, 79–86.

[25] Hisao Katsumi, Wataru Yamada, and Keiichi Ochiai. 2023. Characterizing Generic POI: A Novel Approach for Discovering Tourist Attractions. *J. of Information Processing* 31 (2023), 265–277.

[26] Bart P. Knijnenburg and Martijn C. Willemsen. 2015. *Evaluating Recommender Systems with User Experiments*. Springer, Boston, MA, 309–352.

[27] Yehuda Koren and Robert Bell. 2011. *Advances in Collaborative Filtering*. Springer, Boston, MA, 145–186.

[28] Anna Kovacs-Györi, Alina Ristea, Ronald Kolcsar, Bernd Resch, Alessandro Crivellari, and Thomas Blaschke. 2018. Beyond Spatial Proximity—Classifying Parks and Their Visitors in London Based on Spatiotemporal and Sentiment Analysis of Twitter Data. *ISPRS* 7, 9 (Sept. 2018).

[29] Debolina Kundu and Arvind Kumar Pandey. 2020. World urbanisation: trends and patterns. *Developing national urban policies: Ways forward to green and smart cities* (2020), 13–49.

[30] Urszula Kużelewska. 2020. *Effect of Dataset Size on Efficiency of Collaborative Filtering Recommender Systems with Multi-clustering as a Neighbourhood Identification Strategy*. Springer, 342–354.

[31] Oh-Hyun Kwon, Inho Hong, Jeasurk Yang, Donghee Y. Wohn, Woo-Sung Jung, and Meeyoung Cha. 2021. Urban green space and happiness in developed countries. *EPJ Data Science* 10, 1 (May 2021).

[32] Kate Lachowycz and Andy P. Jones. 2013. Towards a better understanding of the relationship between greenspace and health: Development of a theoretical framework. *Landscape Urban Plan.* 118 (Oct. 2013), 62–69.

[33] Géraud Le Falher, Aristides Gionis, and Michael Mathioudakis. 2015. Where is the Soho of Rome? Measures and algorithms for finding similar neighborhoods in cities. In *ICWSM*, Vol. 9. 228–237.

[34] Xiaojiang Li and Debarchana Ghosh. 2018. Associations between Body Mass Index and Urban "Green" Streetscape in Cleveland, Ohio, USA. *Int. J. Env. Res. Pub. He.* 15, 10 (Oct. 2018), 2186.

[35] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and Shanika Karunasekera. 2018. Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. *Knowl. Inf. Syst.* 54, 2 (Feb. 2018), 375–406.

[36] Qi Liu, Hidayat Ullah, Wanggen Wan, Zhangyou Peng, Li Hou, Sanam Shahla Rizvi, Saqib Ali Haidery, Tong Qu, and A. A. M. Muzahid. 2020. Categorization of Green Spaces for a Sustainable Environment and Smart City Architecture by Utilizing Big Data. *Electronics* 9, 6 (June 2020), 1028.

[37] Jolanda Maas, Robert A Verheij, Peter P Groenewegen, Sjerp de Vries, and Peter Spreeuwenberg. 2006. Green space, urbanity, and health: how strong is the relation? *J. of Epidemiology & Community Health* 60, 7 (July 2006), 587–592.

[38] Madison Milne-Ives, Ching Lam, Caroline De Cock, Michelle Helena Van Velthoven, and Edward Meinert. 2020. Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. *JMIR mHealth and uHealth* 8, 3 (March 2020). https://doi.org/10.2196/17046

[39] Nikola Milojevic-Dupont, Nicolai Hans, Lynn H. Kaack, Marius Zumwald, François Andrieux, Daniel de Barros Soares, Steffen Lohrey, Peter-Paul Pichler, and Felix Creutzig. 2020. Learning from urban form to predict building heights. *PLOS ONE* 15, 12 (Dec. 2020).

[40] A Milvoy and A Roué-Le Gall. 2015. Aménager des espaces de jeux favorables à la santé. *La Santé en action* 444 (Dec. 2015), 38–39.

[41] Daniel F Morse, Sahil Sandhu, Kate Mulligan, Stephanie Tierney, Marie Polley, Bogdan Chiva Giurca, Siân Slade, Sónia Dias, Kamal R Mahtani, Leanne Wells, Huali Wang, Bo Zhao, Cristiano Emanuel Marta De Figueiredo, Jan Joost Meijs, Hae Kweun Nam, Kheng Hock Lee, Carolyn Wallace, Megan Elliott, Juan Manuel Mendive, David Robinson, Miia Palo, Wolfram Herrmann, Rasmus Østergaard Nielsen, and Kerryn Husk. 2022. Global developments in social prescribing. *BMJ Global Health* 7, 5 (2022). https://doi.org/10.1136/bmjgh-2022-008524

[42] Kheng Ng, Angelia Sia, Maxel Ng, Crystal Tan, Hui Chan, Chay Tan, Iris Rawtaer, Lei Feng, Rathi Mahendran, Anis Larbi, Ee Kua, and Roger Ho. 2018. Effects of Horticultural Therapy on Asian Older Adults: A Randomized Controlled Trial. *Int. J. Env. Res. Pub. He.* 15, 8 (Aug. 2018), 1705.

[43] Agnieszka Olszewska-Guizzo, Angelia Sia, Anna Fogel, and Roger Ho. 2020. Can Exposure to Certain Urban Green Spaces Trigger Frontal Alpha Asymmetry in the Brain?—Preliminary Findings from a Passive Task EEG Study. *Int. J. Env. Res. Pub. He.* 17, 2 (Jan. 2020), 394.

[44] Agnieszka Olszewska-Guizzo, Angelia Sia, Anna Fogel, and Roger Ho. 2022. Features of urban green spaces associated with positive emotions, mindfulness and relaxation. *Sci. Rep.* 12, 1 (Nov. 2022).

[45] Halil Özgüner. 2011. Cultural differences in attitudes towards urban parks and green spaces. *Landscape Research* 36, 5 (2011), 599–620.

[46] Matthew J Page et al. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* (March 2021), n71.

[47] Miisa Pietilä, Marjo Neuvonen, Katja Borodulin, Kalevi Korpela, Tuija Sievänen, and Liisa Tyrväinen. 2015. Relationships between exposure to urban green spaces, physical activity and self-rated health. *J. of Outdoor Recreation and Tourism* 10 (July 2015), 44–54.

[48] Luis Pinto, Carla S.S. Ferreira, and Paulo Pereira. 2021. Environmental and socioeconomic factors influencing the use of urban green spaces in Coimbra (Portugal). *Sci. Total Environ.* 792 (Oct. 2021), 148293.

[49] Anna Ponjoan, Jordi Blanch, Lia Alves-Cabratosa, Ruth Martí-Lluch, Marc Comas-Cufí, Marta Cirach, Mark Nieuwenhuijsen, María del Mar Garcia-Gil, and Rafel Ramos. 2022. Impact of residential greenness on myocardial infarction in the population with diabetes: A sex-dependent association? *Environ. Res.* 205 (April 2022), 112449.

[50] Adam Poulston, Mark Stevenson, and Kalina Bontcheva. 2017. Hyperlocal Home Location Identification of Twitter Profiles. In *HT*. ACM, NY, USA, 45–54.

[51] Kostis Pristouris, Harry Nakos, Yannis Stavrakas, Konstantinos I Kotsopoulos, Theofanis Alexandridis, Myrto S Barda, and Konstantinos P Ferentinos. 2021. An integrated system for urban parks touring and management. *Urban Science* 5, 4 (2021), 91.

[52] Daniele Quercia, Joao Paulo Pesce, Virgilio Almeida, and Jon Crowcroft. 2013. Psychological maps 2.0: a web engagement enterprise starting in London. In *WWW*. ACM, NY, USA, 1065–1076.

[53] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *EMNLP*. ACL.

[54] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *UAI*. AUAI Press, 452–461.

[55] Steffen Rendle, Zeno Gantner, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2011. Fast context-aware recommendations with factorization machines. In *SIGIR*. ACM, NY, USA, 635–644.

[56] Raheleh Rostami, Hasanuddin Lamit, Seyed Meysam Khoshnava, and Rasoul Rostami. 2014. The Role of Historical Persian Gardens on the Health Status of Contemporary Urban Residents. *EcoHealth* 11, 3 (May 2014), 308–321.

[57] Susanne Röhr, Francisca S. Rodriez, Rosa Siemensmeyer, Felix Müller, Roman Romero-Ortuno, and Steffi G. Riedel-Heller. 2021. How can urban environments support dementia risk reduction? *Int. J. Geriatr. Psych.* 37, 1 (Oct. 2021).

[58] Diego Saez-Trumper, Daniele Quercia, and Jon Crowcroft. 2012. Ads and the City: Considering Geographic Distance Goes a Long Way. In *RecSys*. ACM, NY, USA, 187–194.

[59] Pablo Sánchez and Alejandro Bellogín. 2022. Point-of-interest Recommender Systems Based on Location-based Social Networks: A Survey from an Experimental Perspective. *Comput. Surveys* (Jan. 2022).

[60] Dieneke Schram-Bijkerk, Piet Otte, Liesbet Dirven, and Anton M. Breure. 2018. Indicators to support healthy urban gardening in urban management. *Sci. Total Environ.* 621 (April 2018), 863–871.

[61] Danielle F. Shanahan, Robert Bush, Kevin J. Gaston, Brenda B. Lin, Julie Dean, Elizabeth Barber, and Richard A. Fuller. 2016. Health Benefits from Nature Experiences Depend on Dose. *Sci. Rep.* 6, 1 (June 2016).

[62] Xiao Ping Song, Daniel R. Richards, and Puay Yok Tan. 2020. Using social media user attributes to understand human–environment interactions at urban parks. *Sci. Rep.* 10, 1 (Jan. 2020).

[63] Bart Thomee, David A. Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: The New Data in Multimedia Research. *Commun. ACM* 59, 2 (Jan. 2016), 64–73.

[64] Caoimhe Twohig-Bennett and Andy Jones. 2018. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environmental research* 166 (2018), 628–637.

[65] Department of Economic United Nations and Social Affairs Sustainable Development. 2015. Transforming our World: the 2030 Agenda for Sustainable Development. , 16301 pages. https://sdgs.un.org/2030agenda

[66] David Vlahov, Nicholas Freudenberg, Fernando Proietti, Danielle Ompad, Andrew Quinn, Vijay Nandi, and Sandro Galea. 2007. Urban as a Determinant of Health. *J. Urban Health* 84, S1 (March 2007), 16–26.

[67] Carole-Jean Wu, Robin Burke, Ed H. Chi, Joseph Konstan, Julian McAuley, Yves Raimond, and Hao Zhang. 2020. Developing a Recommendation Benchmark for MLPerf Training and Inference. https://doi.org/10.48550/ARXIV.2003.07336

[68] Shengbiao Wu, Bin Chen, Chris Webster, Bing Xu, and Peng Gong. 2023. Improved human greenspace exposure equality during 21st century urbanization. *Nat. Commun.* 14, 1 (Oct. 2023).

[69] Hongzhi Yin, Bin Cui, Jing Li, Junjie Yao, and Chen Chen. 2012. Challenging the long tail recommendation. *VLDB Endowment* 5, 9 (May 2012), 896–907.

[70] Hamed Zamani, Markus Schedl, Paul Lamere, and Ching-Wei Chen. 2019. An Analysis of Approaches Taken in the ACM RecSys Challenge 2018 for Automatic Music Playlist Continuation. *TIST* 10, 5, Article 57 (sep 2019), 21 pages.

[71] Chengcheng Zeng, Wei Lin, and Qibing Chen. 2023. Psychophysiological Responses of College Students to Audio-Visual Forest Trail Landscapes. *J. Urban Health* 100, 4 (July 2023), 711–724.

[72] Liqing Zhang, Puay Yok Tan, and Jessica A Diehl. 2017. A conceptual framework for studying urban green spaces effects on health. *J. Urban Ecol.* 3, 1 (Jan. 2017).

[73] Yi-Liang Zhao, Liqiang Nie, Xiangyu Wang, and Tat-Seng Chua. 2014. Personalized recommendations of locally interesting venues to tourists via cross-region community matching. *TIST* 5, 3 (2014), 1–26.

[74] Cynthia Zutter and Ashley Stoltz. 2023. Community gardens and urban agriculture: Healthy environment/healthy citizens. *Int. J. Ment. Health Nu.* (April 2023).