electronics

Article

Personalized Tour Recommendation via Analyzing User Tastes for Travel Distance, Diversity and Popularity

Jongsoo Lee [1], Jung Ah Shin [2], Dong-Kyu Chae [1] and Sang-Chul Lee [3,*]

1 Department of Computer and Software, Hanyang University, Seoul 04763, Korea; leejongsoo@hanyang.ac.kr (J.L.); dongkyu@hanyang.ac.kr (D.-K.C.)
2 The Data Science Institute, Columbia University, New York, NY 10027, USA; js5569@columbia.edu
3 Division of Nanotechnology, Daegu Gyeongbuk Institute of Science & Technology (DGIST), Deagu 42988, Korea
* Correspondence: sangchul.lee@dgist.ac.kr

Abstract: The goal of a tour recommendation is to recommend the best destinations according to the preferences of each tourist. The task of tour recommendation is challenging in that it not only has to consider the ratings, as do existing traditional recommendation problems, but it must also consider the personalization of the unique characteristics, such as diversity, travel distance, and popularity of the travel destination, which previous studies have failed to take into account. In this paper, we propose, for the first time, aspect personalization: we find out how important each user considers the diversity, distance and popularity of a travel destination when choosing where to visit. Then, we provide recommendations on tourist attractions by combining the personalized score for each factor and the predicted score. For the evaluation, we gathered user ratings and metadata of POIs from TripAdvisor and Naver. Experimental results showed that the proposed method had an 82%, 24% and 20% improvement in precision and a 129%, 35% and 22% improvement in recall in terms of top-1, top-2 and top-3 recommendations.

Keywords: tour recommendation; diversification; taste variations; personalization; user taste prediction

check for updates

Citation: Lee, J.; Shin, J.A.; Chae, D.-K.; Lee, S.-C. Personalized Tour Recommendation via Analyzing User Tastes for Travel Distance, Diversity and Popularity. *Electronics* **2022**, *11*, 1120. https://doi.org/10.3390/electronics11071120

Academic Editor: Prasan Kumar Sahoo

Received: 27 February 2022
Accepted: 30 March 2022
Published: 1 April 2022

1. Introduction

Tourism is one of the largest leisure industries. Nearly 1.45 billion people travel and spend USD 1.48 trillion every year [1]. As both private and public transportation improves, the number of new travelers and possible travel destinations has increased. Because of the increase, travelers are required to spend an extensive time making a proper decision. To efficiently review increased destination options, points-of-interest (POIs) data of massive user experience on the Internet can be used [2,3]. With countless amounts of data, recommender systems can assist a user's decision making based on one's personal preference [4,5]. For tour recommendations, tourism-related data collected from social media are expected to be significant for suggesting personalized POIs. The aim of this study is to develop a recommender system that provides suitable POIs to users.

The factors that must be considered when choosing a destination vary from person to person. Some travelers prefer popular destinations that are close to their accommodation, while other travelers may prefer to visit not-well-known, hidden spots. Some travelers may want to visit a variety of venues, such as parks, museums and shopping malls, while others may prefer to visit only certain types of locations. There are thus various factors that go into the POI selection, and it is very important to personalize the recommendations for each user, with those factors taken into consideration.

Existing studies have focused on the nonpersonalized popularity of POI, and the user's time budget, including the total travel time, which considers the distance traveled between travel destinations and the time spent at the destination [6,7]. Regardless of the user's personal preference, if a place is well-known, that place unconditionally has a high popularity score. Therefore, it does not take into account the preference of users who do not consider the popularity of the destination as an important factor. Moreover, personalized importance for other factors, such as tour distance or diversity, are rarely considered [8].

In order to solve the problem, this study focused on the personalization of each element for each user. Each user's personal preference for popularity, distance and diversity was inferred based on the ratings that each user had given to travel destinations in the past. The personalized scores for each element were calculated as follows:

- **Personalized diversity score (*p-Div*)**: This score reflects a user's preference score for each category, and is obtained by counting the categories of the travel destinations each user has visited in the past. If the user repeatedly visits only a specific category, the score of that category is relatively high; when a user visits various categories, the score distribution is even.
- **Personalized popularity score (*p-Pop*)**: This score indicates how much importance a user places in the popularity of the travel destination. This score is obtained based on the average popularity of the travel destinations that the user has visited in the past. This score is higher if the user prefers famous tourist destinations, and lower otherwise. It controls the impact of the popularity of the tourist destination on the final recommendation.
- **Personalized distance score (*p-Dis*)**: This score indicates how much the user considers the travel distance in selecting the POI. It is determined based on the average distance between destinations that the user has visited in the past. The score is high if the average distance is short, and a lower weight is given for longer average distance, thus controlling the impact distance has in recommending a tourist destination based on the distance between the user's estimated location and the destination.

We trained an autoencoder with the ratings left by users in POIs to predict ratings for POIs that users have not yet visited. Users' more recent ratings provided a greater weight to the model's training. Then, the top N POIs were recommended after deriving the final score, by summing the personalized scores for each aspect, as described above. The proposed model was evaluated based on data from one of the popular travel destinations in South Korea, Jeju Island. Specifically, we collected 156 POIs located on Jeju Island on TripAdvisor, 29,020 ratings left by 7718 users who visited the venue, 109 POIs located on Jeju Island in Naver, and 270,806 ratings left by 109,754 users who visited the Island. As a result of experiments based on these data, the highest recommendation accuracy was observed when each factor was personalized and reflected in the recommendation as we suggested.

The main contributions of this paper can be summarized as follows:

- We investigated a way to quantify a user's personal preference for each travel-related aspect (diversity, popularity and distance) based on their tour history.
- We proposed a novel tour recommendation method that is able to consider each user's personalized taste for various aspects of POIs as well as their predicted ratings on them.
- We conducted extensive experiments to evaluate the proposed method. The results show that our idea of considering user tastes for various aspects is really effective in recommending potential POIs, which makes our method outperform several baseline methods.

The remainder of the paper is structured as follows: Section 2 describes related studies. In Section 3, we introduce our data and define some notations. Section 4 describes the proposed method in detail. Section 5 details the experimental environment, and Section 6 reports the experimental results. Finally, Section 7 summarizes the conclusions and introduces future research topics.

2. Related Work

As recommendation systems have become popular, tour recommendations have also become one of the important research areas [6–12]. Most of the studies produced recommendations based on social media data, such as Flickr, Yelp and Foursquare [13]. In the case of Flickr, information about images, locations and times taken in geotagged photos were mainly used [6,7,11,14], and for Yelp and Foursquare, the ratings on POIs were mainly used [8,12]. In one study [11], the users' preferences were identified by measuring the time duration at the POI based on the location and time information included in the users' photographs. In studies [6,7], the places visited were identified using location data where the photos were taken, and users' preferences for each travel destination category were identified based on the category distribution of those places. Ref. [8] used the ratings given by users to POIs to understand users' preferences.

For a successful travel destination recommendation, it is important to understand how important popularity, diversity and distance are for each user. However, most existing studies miss one or more of the aforementioned factors. Table 1 summarizes the factors that related studies consider as part of personalization, and those that do not. For popularity, most studies give high scores to famous travel destinations regardless of whether the users value popularity as an important factor [6,7,9,11,12]. In the case of the distance to the destination, most studies only used this for calculating the time budget. For example, in [7,11], the user's movement speed was predefined (4 km/h or 5 km/h) and the destination was selected so that the time taken to reach the location did not exceed the user's time budget. In [6], the authors would only recommend places based on setting the distance budget between POIs and making sure the distance between POIs did not exceed the distance budget. However, to the best of our knowledge, there have not been any studies reflecting the users' preferences towards distance factor, whether they preferred closeness to a POI or not.

Table 1. Comparisons of related work and our method. X: does not reflect that aspect in recommendations; △ : reflects the aspect, but does not consider personalization of that aspect; O: personalizes and reflects that aspect into recommendations.

Methods	Popularity Personalization	Diversity Personalization	Distance Personalization
Based on user interests and visit durations [11]	△	○	△
POI availability and uncertain traveling time [8]	X	X	△
User interests from geotagged photos [6]	△	X	△
Based on queuing time [7]	△	○	△
Learning points and routes [9]	△	○	△
Aurigo [12]	X	○	○
Ours	○	○	○

Of the existing research, only Aurigo [12] personalized both distance and diversity aspects and reflected both aspects in their recommendations. However, since Aurigo received direct input about the distance criterion and preference of category from users, personalization was possible. This method can therefore only be used in an environment in which direct interaction with the user is possible. We propose a method to personalize all of the popularity, distance and diversity metrics, and reflect them in recommendations without requiring an interactive system.

In the context of collaborative filtering, an autoencoder-based rating prediction and recommendation has been studied [15–17], which is also related to this work. An autoencoder is a deep neural architecture trained to produce its output as similar with the input as possible. In the recommender systems area, the autoencoder is used to reconstruct a dense vector with predictions on the missing entries by feeding a sparse vector [15,16]. The early work [15] proposed user-specific autoencoder for recommendation, while AutoRec [16] used just one autoencoder that was trained based on the entire dataset. There are two variants of AutoRec depending on input types: user-based AutoRec uses user rating vectors as inputs and item-based AutoRec uses item rating vectors as inputs. HybridAE [17] suggested a hybrid recommender system based on an autoencoder, which incorporated user-item rating matrix and contents information for the solution to the lack of information.

3. Preliminaries

3.1. Data Description

We collected data from Jeju Island, one of the most famous tourist cities in Korea. Information and reviews on travel destinations in Jeju were collected from TripAdvisor [18] and Naver. We crawled 156 registered destinations and collected 7718 users' 29,020 reviews from TripAdvisor. Based on 156 TripAdvisor destinations, we collected 109,754 users' 270,806 reviews posted on Naver Map. The dataset includes the anonymous user ID, rating and date. Ratings were scored on scales of 1 to 5 and 0.5 to 5, respectively, for TripAdvisor and Naver. Table 2 shows the statistics of the dataset we collected.

Table 2. Data statistics.

Dataset	#Users	#POIs	#Ratings	#Categories
TripAdvisor	7718	156	29,020	10
Naver	109,754	109	270,806	8

3.2. Preliminaries and Notations

The notations used in our work are summarized in Table 3. We define a list of (user, POI, category), tour history, rating matrix, and diversity matrix for each user as follows:

- (User, POI, category) list: We have a set of m users, a set of n POIs, and a set of i categories, respectively, defined as: $\mathcal{U} = \{u_1, u_2, \ldots, u_m\}$, $\mathcal{P} = \{p_1, p_2, \ldots, p_n\}$, $\mathcal{C} = \{c_1, c_2, \ldots, c_i\}$. Each POI p_n is associated with just one category.
- Tour history: The tour history of each user is defined as $\mathcal{H}_u = ((p_1, r_{up_1}, t^u_{p_1}, n_{p_1}, e_{p_1}, c_{p_1}), \ldots, (p_n, r_{up_n}, t^u_{p_n}, n_{p_n}, e_{p_n}, c_{p_n}))$. Here, the components of each tuple are POI p_n visited by user u, the rating r_{up_n} left on p_n, the date the rating was given $t^u_{p_n}$, the latitude n_{p_n} and longitude e_{p_n} of the POI and the POI category c_{p_n}.
- Rating matrix: We aggregate the rating data included in the tour history of all users and combine the ratings left by all users on the POIs into a single sparse matrix $\mathbb{R} = (r_{up})_{m \times i}$. Element r_{up} has a rating if user u left a rating on POI p.
- Diversity matrix: We aggregate the category information included in each user's tour history by calculating the number of visits for each category by all users and creating a sparse matrix $\mathbb{D} = (d_{uc})_{m \times n}$. The number of times user u visited category c is represented as d_{uc}, computed by:

$$d_{uc} = \sum_{u_x \in \mathcal{U}} \sum_{p \in \mathcal{H}_u} \delta(u = u_x) \cdot \delta(c = c_p) \tag{1}$$

where $\delta(u = u_x)$ is an indicator function that returns 1 when u_x and u are equal, and returns 0 when they are not. If u has never visited c, a value of 0, meaning empty, is assigned to d_{uc}.

Table 3. Notations.

Symbol	Description
$\mathcal{U}$	a set of users
u	a user
$\mathcal{P}$	a set of POIs
p	a POI
$\mathcal{C}$	a set of categories
c	a category
$\mathcal{H}_u$	tour history of user u
n_p	latitude of POI p
e_p	longitude of POI p
c_p	category of POI p
r_{up}	rating of POI p by user u
t_p^u	date of rating left by user u at POI p
d_{uc}	diversity score of category c by user u
$\mathbb{R}$	rating matrix
$\mathbb{D}$	diversity matrix

4. Method

4.1. Overview

Our main aim was to identify each user's personalized preferences for popularity, diversity and distance, and recommend travel destinations based on these data. Each personalized score was defined as *p-Div*, *p-Pop* and *p-Dis*, and the way in which these values were calculated are explained in detail in the next sections. Then, we calculated the final score $\mathcal{S}(u, p)$ for the POI p of user u by obtaining a normalized weighted sum of (1) predicted the rating of u on p, (2) popularity of p and (3) distance from u to p, where each of the personalized scores *p-Div*, *p-Pop* and *p-Dis* were used as weights. Finally, we suggested the user's top N destinations based on the highest $\mathcal{S}(u, p)$. The conceptual overview of our recommendation system is illustrated in Figure 1.

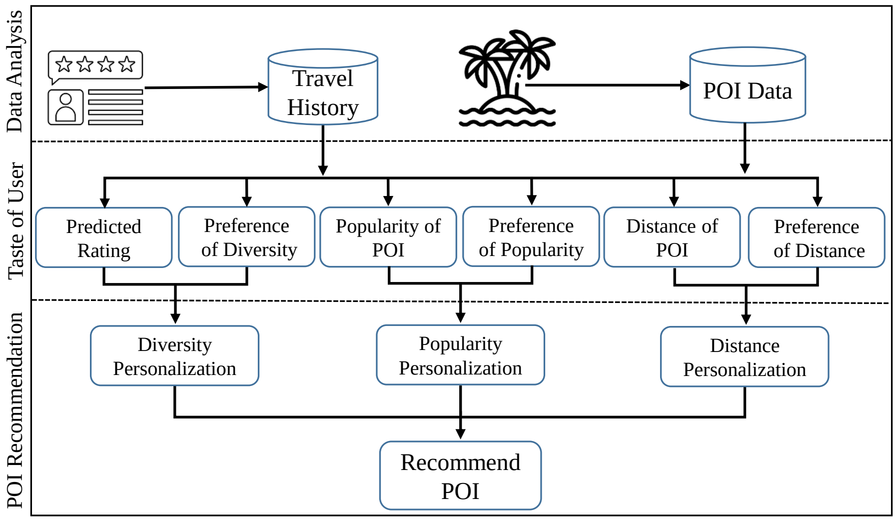

Figure 1. System overview.

4.2. Diversity Personalization

This section describes the adaptation of personal preference for diversity of POIs in the POI recommendation. We defined a value *p-Div*(u, p): it is high for a POI p from a specific category if u only visited a certain type of destination; conversely, if u visited diverse types of destinations, the value *p-Div*(u, p) would be evenly high for various categories. Then, to reflect the characteristic in the prediction, we also predicted each user's rating for each POI and multiplied the predicted rating by *p-Div*(u, p) as a weight.

We used an autoencoder to predict the user's rating and the value of *p-Div*. Each column vector of the sparse matrix $\mathbb{D}$ composed of the number of visits by category was used as an input–output to train the autoencoder, and the dense matrix $\hat{\mathbb{D}}$ was derived by using the aggregated reconstructed output from the trained autoencoder. The loss function of the autoencoder was calculated as follows:

$$\mathcal{L}(\mathbb{D}) = \sqrt{(\frac{1}{m}) \sum_{u=1}^{m} (\mathbf{d}_u - \hat{\mathbf{d}_u})^2} \tag{2}$$

where $\mathbf{d}_u$ is u's (sparse) category preference vector and $\hat{\mathbf{d}_u}$ is the predicted (dense) category preference vector. We defined each element d_{uc} of $\hat{\mathbb{D}}$ as the value of $p\text{-}Div(u,p)$ for u's category c. We consider that u views diversity to be an important factor in POI selection if the distribution of a user's $p\text{-}Div(u,p)$ is evenly distributed across all categories. If $p\text{-}Div(u,p)$ is only high for certain categories, it can be regarded as the user preferring a specific category rather than a diversity of categories.

Next, we predicted the ratings for unvisited POIs. We trained the autoencoder using each column vector of $\mathbb{R}$. Then, using the trained autoencoder, we derived a dense matrix $\hat{\mathbb{R}}$ containing predicted ratings. The loss function used for training was as follows:

$$\mathcal{L}(\mathbb{R}) = \sqrt{(\frac{1}{\sum_{u=1}^{m} \delta(\mathbf{r}_u \neq 0)}) \sum_{u=1}^{m} (\mathbf{r}_u - \hat{\mathbf{r}_u})^2 \cdot \delta(\mathbf{r}_u \neq 0) \cdot \mathbf{t}_u} \tag{3}$$

where $\delta(r_{up} \neq 0)$ is an indicator function that returns 1 if u leaves a rating on p, otherwise, 0. By doing so, we limited the loss only to observed elements in the vectors. Note that, in Equation (2), we did not use the term $\delta(r_{up} \neq)$ so that the zeros included in $\mathbb{D}$ could also be considered in the loss function. This is because missing entries in $\mathbb{D}$ are meaningful in that they convey information that a specific category has not been visited. However, we chose to use the indicator function for learning $\mathbb{R}$ since it is an extremely sparse matrix that includes too many zeros, which might dominate the small number of observed ratings [19].

In addition, we defined a weight vector $\mathbf{t}_u$ that has a larger value if the ratings left by a user was more recent. This was done so that the more recent ratings would have a greater influence on model training. Each element $t_{u,p}$ of this vector was computed as follows:

$$t_{u,p} = 0.5 + \frac{t_p^u - min(t^u)}{max(t^u) - min(t^u)} \times 0.5 \tag{4}$$

where $max(t^u)$ and $min(t^u)$ are the maximum and minimum values of the time at which u left a rating and are used for normalization.

Finally, we applied the value we acquired by multiplying each element $\hat{r_{up}}$ of $\hat{\mathbb{R}}$ with u's $p\text{-}Div(u,p)$ from $\hat{\mathbb{D}}$'s (u,c) element for the final recommendation.

4.3. Popularity Personalization

In this section, we explain how the popularity of a POI is personalized and considered in the recommendations. First, the absolute popularity score of each POI was obtained, and then the $p\text{-}Pop$ score, which indicates how important each user considers the popularity aspect to be, was calculated. Then, the two values were simply multiplied and reflected in the final recommendation.

The nonpersonalized popularity score of a POI p, denoted as $Pop(p)$, was calculated using the number of nonduplicate visitors to p and the average of the ratings left by them. The specific formula was as follows:

$$Pop(p) = \frac{r_p}{Cnt(p)} + \frac{Cnt(p)}{\sum_{p_y \in \mathcal{P}} Cnt(p_y)} \times 5 \tag{5}$$

where each r_p and $Cnt(p)$ was computed by:

$$r_p = \sum_{u \in \mathcal{U}} \sum_{p_y \in \mathcal{H}_u} \delta(p = p_y) \cdot r_{up} \tag{6}$$

$$\text{Cnt}(p) = \sum_{u \in \mathcal{U}} \sum_{p_y \in \mathcal{H}_u} \delta(p = p_y) \tag{7}$$

Equation (6) calculates the sum of the ratings left by users for p. Equation (7) counts the number of ratings that users have left in p. Since the average of the ratings, which is the first term of Equation (5), can have a value between 1 and 5, the second term is multiplied by 5 to match the ranges of the two terms.

Next, we calculated the *p-Pop* score, which implies how much the user considers popularity when selecting a POI. We first plotted the average popularity distribution of POIs visited by each user. The results are shown in Figure 2, in which the left graph displays the average popularity distribution of POIs visited by TripAdvisor users, and the graph on the right is that of Naver users. Using these graphs, we can see that some users only visit famous POIs, whereas others do not. To capture this aspect, we defined the personalized popularity weight *p-Pop* for user u as the average popularity of POIs visited by u as follows:

$$p\text{-}Pop(u) = \frac{Pop(u)}{Cnt(u)} \tag{8}$$

where each $Pop(u)$ and $Cnt(u)$ was computed by:

$$Pop(u) = \sum_{u_x \in \mathcal{U}} \sum_{p_y \in \mathcal{H}_u} \delta(u = u_x) \cdot pop(p_y) \tag{9}$$

$$Cnt(u) = \sum_{u_x \in \mathcal{U}} \sum_{p_y \in \mathcal{H}_u} \delta(u = u_x) \tag{10}$$

Here, Equation (9) is the sum of the popularity of POIs visited by u, and Equation (10) counts the number of POIs visited by u. The higher the value of $p\text{-}Pop(u)$, the more important u values popularity in a POI selection. Finally, we reflected the value obtained by multiplying the popularity value $Pop(p)$ of a POI by $p\text{-}Pop(u)$ in the final recommendation.

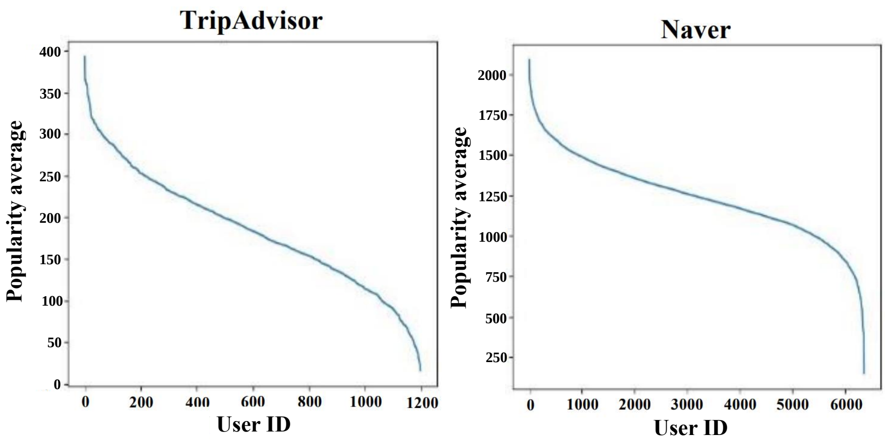

Figure 2. Average of popularity values for each user.

4.4. Distance Personalization

We took into account users' personal preferences on whether or not they preferred short distances to travel destinations when providing recommendations. We first assumed that the center of the coordinates of the POIs visited by user u in the past was the starting point for u. Then, the nonpersonalized distance score between each POI not visited by u was calculated. Next, we derived the *p-Dis* weight, which indicates how important each user considers the distance aspect to be. The result of multiplying the two values was reflected in the final recommendation.

First, the nonpersonalized distance score, denoted as $Dis(u,p)$, from the starting point of u to each nonvisited POI p was calculated as follows:

$$Dis(u,p) = max(\mathcal{H}(u, p_y \in \mathcal{P})) - \mathcal{H}(u,p) \tag{11}$$

where $\mathcal{H}(u,p)$ is a function that calculates the distance from u to p using the Haversine Formula (the Haversine formula determines the great-circle distance between two points on a sphere given their longitudes and latitudes [20]). In order to give a higher score for a closer distance, we calculated "the distance to the further POI that u could go to" minus "the distance to each p". As a result, the closer the distance between u and p, the higher the value of $Dis(u,p)$.

Next, we computed *p-Dis*. We first plotted the average distance distribution of POIs visited by each visitor. The result is shown in Figure 3, in which the graph on the left shows the average distance distribution of POIs visited by TripAdvisor users and the graph on the right shows the ones visited by Naver users. From the graph, we can confirm that some users only visit POIs nearby while others do not. To capture this aspect, we calculated the personalized distance weight *p-Dis* of u by using the average of the POIs visited by u as follows:

$$p\text{-}Dis(u) = max(avg(u_x \in \mathcal{U})) - avg(u) \tag{12}$$

where $avg(u)$ indicates the average distance between POIs visited by u. The higher the value of $p\text{-}Dis(u)$, the higher the distance is considered as an important factor in u's POI selection. Finally, the value obtained by multiplying the distance value $Dis(u,p)$ of the POI by $p\text{-}Dis(u)$ was reflected in the final recommendation.

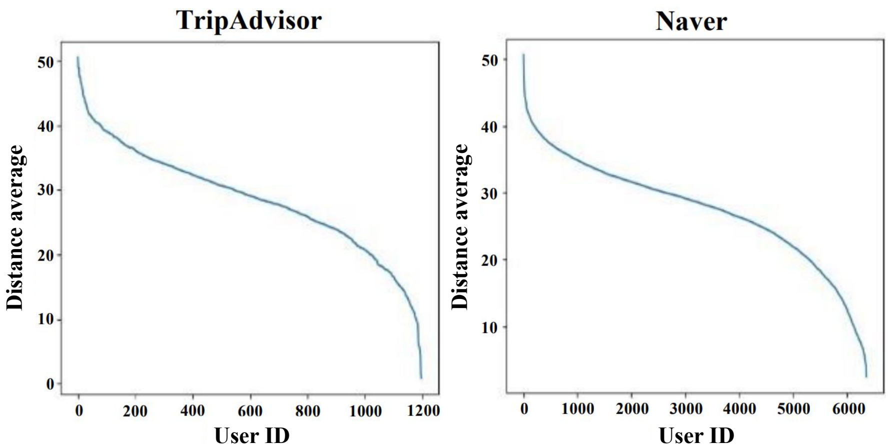

Figure 3. Average of distance (km) to POIs visited for each user.

4.5. Final Score Computation

With the aggregation of the scores obtained from the previous sections, we calculated the final score $\mathcal{S}(u,p)$ for the POI p for user u. We first normalized the previously acquired nonpersonalized values $Pop(p)$ and $Dis(u,p)$ into a scale from 0 to 5, in line with the rating scale. Then, the final score $\mathcal{S}(u,p)$ was calculated as the weighted sum of the scores for each aspect as follows:

$$\mathcal{S}(u,p) = \alpha \cdot \hat{r_{up}} \cdot p\text{-}Div(u,p) + \beta \cdot Pop(p) \cdot p\text{-}Pop(u) + \gamma \cdot Dis(u,p) \cdot p\text{-}Dis(u) \tag{13}$$

where α, β, γ are hyperparameters that control the importance of diversity, popularity and distance aspects, respectively. We found the optimal values for α, β, γ through a grid search.

5. Experimental Settings

5.1. Dataset

As mentioned in Section 3.1, we conducted experiments using data from TripAdvisor and Naver. Of the 29,020 reviews collected from TripAdvisor and 270,806 reviews collected from Naver, we removed reviews that were either:

- Reviews with missing ratings or user IDs.
- Reviews left by users with less than five tour histories; the user was also excluded from the user list.
- Reviews on a POI with fewer than 10 reviews; the POI was also excluded from the POI list.

Among the remaining reviews, we split, respectively, the past 80% and the recent 20% as training and test data, according to the time the reviews were written. Since the sentiment of the user was already expressed in the review, we did not analyze the review text itself.

We plotted the distribution of the number of visitors by POI and category. The results are shown in Figures 4 and 5. The left graph shows the data from TripAdvisor and the right shows the distribution from the Naver data. In all cases, we could observe a typical power law distribution. Details on each category ID are shown in Table 4.

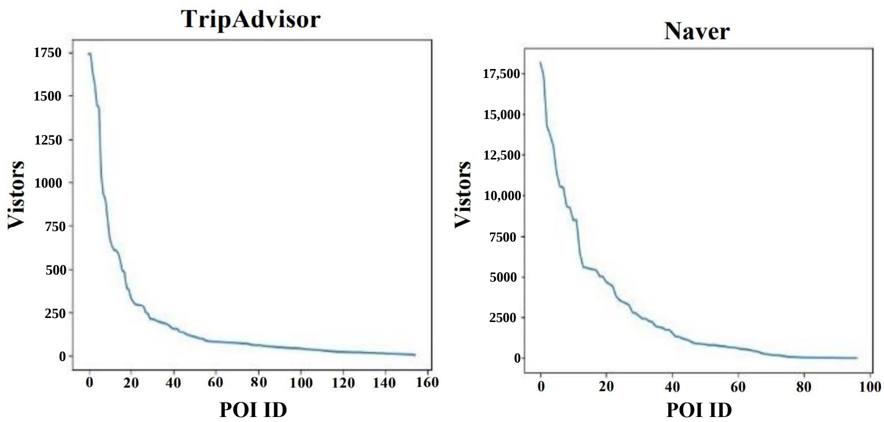

Figure 4. The number of visitors for each POI.

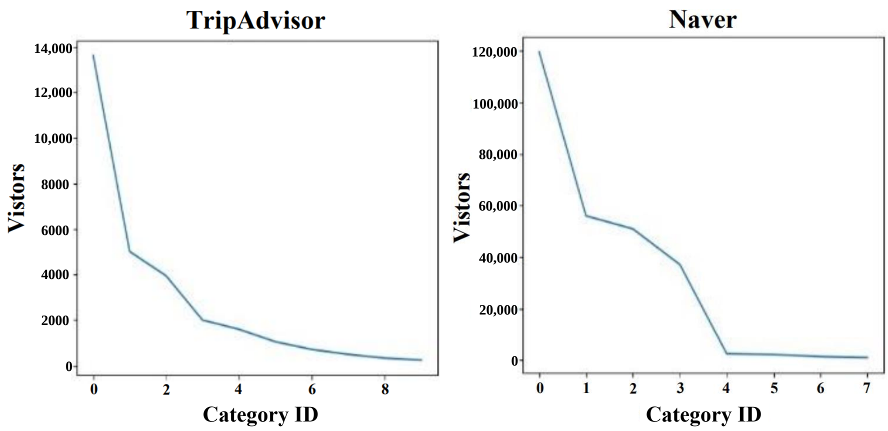

Figure 5. The number of visitors for each category.

Table 4. Categories of POIs.

Category Id	Naver	TripAdvisor
0	theme park	landscape
1	museum	theme park
2	landscape	eco park
3	eco park	museum
4	art museum	shopping
5	shopping	scenic drive
6	activities	activities
7	structure	structure
8		art museum
9		place of worship

5.2. Evaluation Metrics

The accuracy of a recommender system means the ratio of the recommended items appear in the ground truth. In the literature, various metrics have been developed to measure the recommendation accuracy. Among many others, we adopted *precision*, *recall*, *normalized discounted cumulative gain* (*NDCG*) and *mean reciprocal rank* (*MRR*) [21]. Precision and recall are the metrics that quantify how many times a model "hits" the ground truth items. Unlike the previous evaluation metrics, *NDCG* and *MRR* consider the rank of the correct item in the recommended list provided to users.

Formally, for a user u, let Rec_u denote a ranked list of N items recommended to u by an algorithm and $Test_u$ be a set of ground truth items in a test data. In order to evaluate each Rec_u, the four metrics are computed as:

$$Precision_u@N = \frac{|Test_u \bigcap Rec_u|}{|Rec_u|} \tag{14}$$

$$Recall_u@N = \frac{|Test_u \bigcap Rec_u|}{|Test_u|} \tag{15}$$

$$DCG_u@N = \sum_{k=1}^{N} \frac{2^{y_k} - 1}{\log_2(k+1)} \tag{16}$$

$$MRR_u@N = \frac{1}{rank_{first}(u)} \tag{17}$$

where $DCG_u@N$ in Equation (16) denotes the discounted cumulative gain for each user and y_k stands for the relevance score of the kth ranked item in Rec_u to user u ($y_k = 1$ if the item is correct and 0 otherwise). Then, the normalized DCG ($NDCG_u@N$) is computed by dividing $DCG_u@N$ with a DCG obtained by an ideal ranking algorithm. $rank_{first}(u)$ in Equation (17) indicates the ranked position of the first correct item among those in Rec_u [22].

5.3. Implementation Details

For the autoencoders in our framework, we used a SELU function and an identity function as the activation functions of the hidden layers and the output layer, respectively. We used Xavier's network initialization approach [23]. We set the minibatch size to 64, the learning rate to 0.001 and the dropout rate to 0.5. Our autoencoder for the rating prediction had the structure $input \to 90 \to 70 \to 90 \to output$. The autoencoder for our category diversity prediction had a relatively simple structure of $input \to 5 \to output$.

6. Results and Analysis

We first plotted a loss curve to check whether the autoencoder was properly trained on the diversity matrix $\mathbb{D}$ and the rating matrix $\mathbb{R}$. The results are shown in Figures 6 and 7. Figure 6 is the loss curve from the autoencoder trained on $\mathbb{R}$, and Figure 7 is the loss curve from the autoencoder trained on $\mathbb{D}$. We observed that all curves converged at appropriate epochs.

Next, we measured the accuracy of the proposed method. The experimental results are summarized in Figure 8 and Tables 5–7. Each method listed in the Algorithm column of each table is a baseline for comparison with the proposed method. First, the above four baselines are random recommendations and algorithms that consider only one of the three aspects. For example, the random method randomly recommends N of the unvisited POIs. The popularity baseline method recommends the top N POIs based on personalized popularity and the rating method recommends the top N POIs based on the rating predictions of the autoencoder. The distance baseline recommends the top N POIs based on the personalized distance score. The next seven baselines are methods that consider more than one aspect. For example, Popularity + Rating adds up only the popularity and rating scores. Popularity + Rating + Diversity, is a recommendation method that considers popularity, rating and diversity. In the case of diversity, since POIs of the same category have the same value, it is not appropriate to consider the diversity score alone, so we ensured that diversity was considered along with the rating. The proposed method considers all aspects.

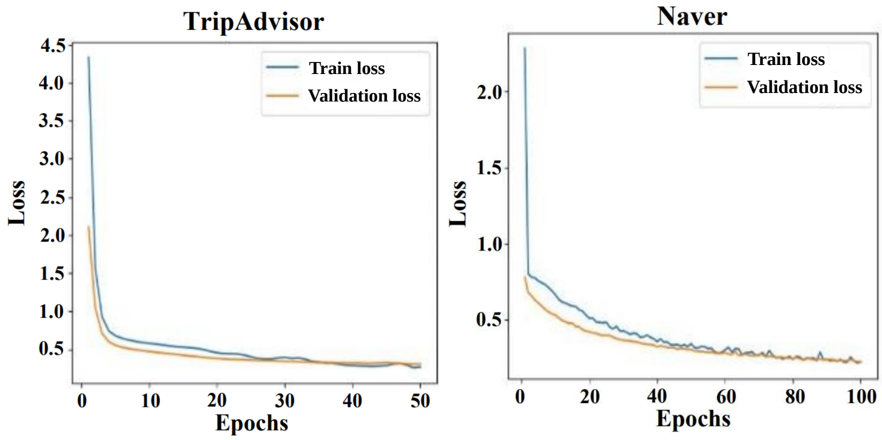

Figure 6. Loss curve (rating prediction).

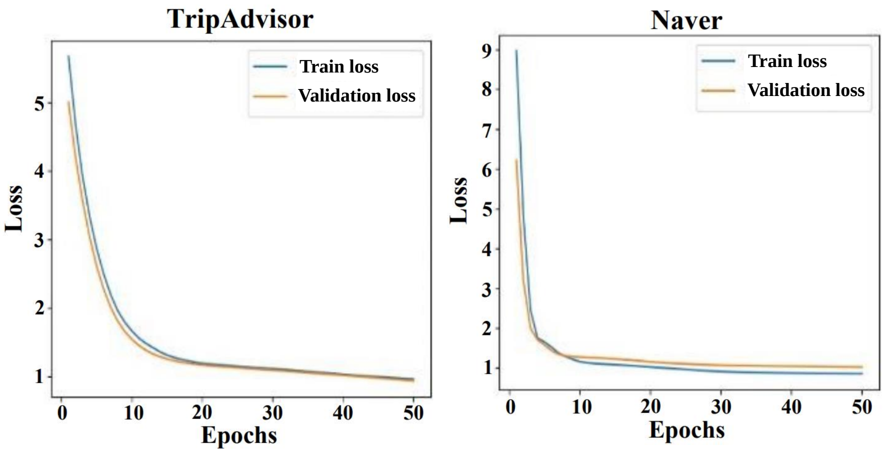

Figure 7. Loss curve (diversity prediction).

We confirmed that the proposed method showed higher accuracy than all the baseline algorithms for both datasets, and for all cases of top-1, top-2, and top-3 recommendations. Among the baselines, all methods considering popularity performed relatively well, and rating, distance and rating+diversity showed relatively low accuracy when used alone. However, since the performance of the proposed method that takes into all of these factors into account was the highest, it can be interpreted that each factor was combined to produce a positive synergistic effect. The proposed method improved by 82%, 24% and 10% the top-1, top-2, and top-3 recommendations compared to the popularity recommendation in the Naver data, respectively, and 34%, 17% and 20% improvement in the TripAdvisor data, respectively. In terms of recall, the recommendation accuracy was improved by 129%, 35% and 10% in the Naver data, and 39%, 17% and 22% in the TripAdvisor data.

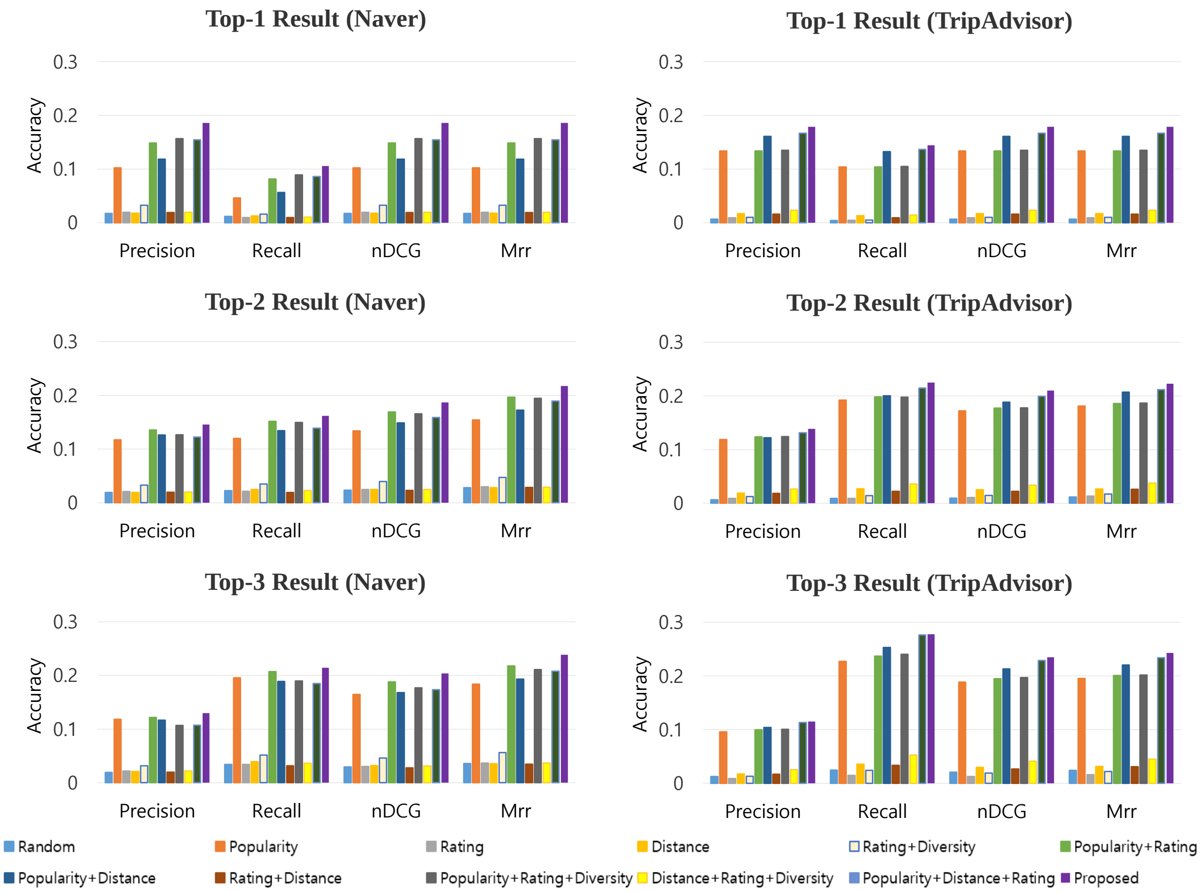

Figure 8. Accuracy comparisons.

Table 5. Top 1 recommendation accuracy comparison.

Dataset	Algorithm	Precision	Recall	nDCG	MRR	Dataset	Algorithm	Precision	Recall	nDCG	MRR
	Random	0.0174	0.012	0.0174	0.0174		Random	0.0067	0.004	0.0067	0.0067
	Popularity	0.1021	0.0462	0.1021	0.1021		Popularity	0.1336	0.1038	0.1336	0.1336
	Rating	0.0195	0.0095	0.0195	0.0195		Rating	0.0092	0.0044	0.0092	0.0092
	Diversity	0.0327	0.0159	0.0327	0.0327		Diversity	0.01	0.0048	0.01	0.01
	Distance	0.0176	0.0125	0.0176	0.0176		Distance	0.0167	0.0131	0.0167	0.0167
Naver	Popularity+Rating	0.1487	0.0815	0.1487	0.1487	TripAdvisor	Popularity+Rating	0.1336	0.1038	0.1336	0.1336
	Popularity+Distance	0.1183	0.0564	0.1183	0.1183		Popularity+Distance	0.1611	0.1323	0.1611	0.1611
	Rating+Distance	0.019	0.0096	0.019	0.019		Rating+Distance	0.0159	0.0091	0.0159	0.0159
	Popularity+Diversity	0.1567	0.0890	0.1567	0.1567		Popularity+Diversity	0.1352	0.1047	0.1352	0.1352
	Distance+Diversity	0.0196	0.0104	0.0196	0.0196		Distance+Diversity	0.0234	0.0143	0.0234	0.0234
	Popularity+Distance+Rating	0.1544	0.0862	0.1544	0.1544		Popularity+Distance+Rating	0.1669	0.1367	0.1669	0.1669
	Proposed	**0.1858**	**0.1056**	**0.1858**	**0.1858**		**Proposed**	**0.1786**	**0.1442**	**0.1786**	**0.1786**

Table 6. Top 2 recommendation accuracy comparison.

Dataset	Algorithm	Precision	Recall	nDCG	MRR	Dataset	Algorithm	Precision	Recall	nDCG	MRR
Naver	Random	0.019	0.0225	0.0233	0.0281	TripAdvisor	Random	0.0067	0.0089	0.0096	0.0117
	Popularity	0.1172	0.1198	0.134	0.1542		Popularity	0.1189	0.192	0.1724	0.1811
	Rating	0.0208	0.0212	0.0245	0.0297		Rating	0.0092	0.0092	0.011	0.0134
	Diversity	0.0327	0.0347	0.0394	0.0474		Diversity	0.0125	0.0144	0.0148	0.0171
	Distance	0.0193	0.0247	0.0246	0.028		Distance	0.0188	0.0269	0.0253	0.0267
	Popularity+Rating	0.1355	0.1519	0.1691	0.1968		Popularity+Rating	0.124	0.1982	0.1773	0.1857
	Popularity+Distance	0.1264	0.1344	0.1486	0.1724		Popularity+Distance	0.1223	0.2002	0.1884	0.207
	Rating+Distance	0.0195	0.0191	0.023	0.0284		Rating+Distance	0.0184	0.0222	0.0223	0.0263
	Popularity+Diversity	0.1267	0.1496	0.1659	0.1948		Popularity+Diversity	0.1244	0.1978	0.1779	0.1866
	Distance+Diversity	0.0201	0.0226	0.0248	0.0291		Distance+Diversity	0.0267	0.0357	0.0339	0.0376
	Popularity+Distance+Rating	0.1224	0.1388	0.1589	0.1893		Popularity+Distance+Rating	0.1315	0.2148	0.1992	0.2116
	Proposed	**0.1457**	**0.1620**	**0.1868**	**0.2177**		**Proposed**	**0.1386**	**0.225**	**0.2099**	**0.2229**

Table 7. Top 3 recommendation accuracy comparison.

Dataset	Algorithm	Precision	Recall	nDCG	MRR	Dataset	Algorithm	Precision	Recall	nDCG	MRR
Naver	Random	0.0194	0.0341	0.0296	0.0357	TripAdvisor	Random	0.0128	0.0244	0.0209	0.0243
	Popularity	0.1183	0.1954	0.1645	0.1837		Popularity	0.0957	0.2273	0.1887	0.1953
	Rating	0.0218	0.0344	0.0302	0.0368		Rating	0.0089	0.0146	0.0127	0.0161
	Diversity	0.0316	0.0515	0.0459	0.0562		Diversity	0.0131	0.0241	0.019	0.0218
	Distance	0.0209	0.0393	0.0319	0.0356		Distance	0.0175	0.0358	0.0296	0.0314
	Popularity+Rating	0.1217	0.207	0.1879	0.2179		Popularity+Rating	0.0996	0.2371	0.1948	0.2008
	Popularity+Distance	0.1168	0.189	0.1679	0.1931		Popularity+Distance	0.1041	0.253	0.2132	0.2205
	Rating+Distance	0.0199	0.0315	0.0279	0.0348		Rating+Distance	0.017	0.0333	0.0266	0.031
	Popularity+Diversity	0.1068	0.1898	0.1770	0.2110		Popularity+Diversity	0.1007	0.2403	0.1968	0.2016
	Distance+Diversity	0.0219	0.0365	0.0313	0.0366		Distance+Diversity	0.0253	0.0527	0.0413	0.0451
	Popularity+Distance+Rating	0.1074	0.1848	0.1734	0.2078		Popularity+Distance+Rating	0.1132	0.2765	0.2288	0.2341
	Proposed	**0.1296**	**0.2144**	**0.2036**	0.2383		**Proposed**	**0.1152**	**0.2779**	**0.2348**	**0.2429**

7. Conclusions and Future Work

In this paper, we proposed a tour recommendation method, where the main idea was to personalize the user's preference for each element. We first built a rating matrix and a diversity matrix for training an autoencoder. Then, a personalized diversity score was derived based on the reconstructed outputs from the autoencoder. Then, the popularity and distance scores of the POI were obtained, as were the degree to which users considered popularity and distance to be important factors in the POI selection, multiplied as weights, and reflected in the final recommendation. The proposed method was evaluated using TripAdvisor and Naver data and showed higher recommendation accuracy than other baselines in all cases.

We believe that there are many other aspects that could be considered in travel destination recommendation aside from the diversity, popularity, and distance factors we considered. In our future work, the weather conditions and the peak season period of each POI will also be considered. Based on the individual POI recommendations, we will work on travel routes recommendations. We also plan to research travel recommendations based on user composition (couples, families, friends and individual) to provide more precise recommendation than current general recommendation.

Author Contributions: Conceptualization, J.A.S. and D.-K.C.; methodology, J.L. and D.-K.C.; software, J.A.S.; validation, J.L. and D.-K.C.; formal analysis, D.-K.C.; investigation, J.L.; resources, J.A.S.; data curation, J.L.; writing—original draft preparation, J.L. and D.-K.C.; writing—review and editing, S.-C.L. and J.A.S.; visualization, J.L. and D.-K.C.; supervision, S.-C.L.; project administration, S.-C.L.; funding acquisition, S.-C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partly supported by (1) the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (no. 2020-0-01373, Artificial Intelligence Graduate School Program (Hanyang University)), (2) the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (no. 2021R1A2C1094863), and (3) the DGIST R&D program of the Ministry of Science and ICT of KOREA (22-IT-10-03).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. UNWTO Tourism Highlights (2020), Edition 2020. Available online: https://www.e-unwto.org/doi/pdf/10.18111/9789284421152 (accessed on 9 February 2022).
2. Ge, Y.; Liu, Q.; Xiong, H.; Tuzhilin, A.; Chen, J. Cost-aware travel tour recommendation. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 983–991.
3. Xiang, Z.; Wang, D.; O'Leary, J.T.; Fesenmaier, D.R. Adapting to the internet: trends in travelers' use of the web for trip planning. *J. Travel Res.* **2015**, *54*, 511–527. [CrossRef]
4. Resnick, P.; Iacovou, N.; Suchak, M.; Bergstrom, P.; Riedl, J. Grouplens: An open architecture for collaborative filtering of netnews. In Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, Chapel Hill, NC, USA, 22–26 October 1997; pp. 175–186.
5. Ricci, F.; Rokach, L.; Shapira, B. *Introduction to Recommender Systems Handbook*, 3rd ed.; Springer: Boston, MA, USA, 2011; pp. 1–35.
6. Lim, K.H. Recommending tours and places-of-interest based on user interests from geo-tagged photos. In Proceedings of the 2015 ACM SIGMOD on PhD Symposium, Melbourne, Australia, 31 May 2015; pp. 33–38.
7. Lim, K.H.; Chan, J.; Karunasekera, S.; Leckie, C. Personalized itinerary recommendation with queuing time awareness. In Proceedings of the 40th international ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Japan, 7–11 August 2017; pp. 325–334.
8. Zhang, C.; Liang, H.; Wang, K.; Sun, J.D. Personalized trip recommendation with poi availability and uncertain traveling time. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, 18–23 October 2015; pp. 911–920.
9. Chen, D.; Ong, C.S.; Xie, L. Learning points and routes to recommend trajectories. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, Indianapolis, IN, USA, 24–28 October 2016; pp. 2227–2232.
10. He, J.; Qi, J.; Ramamohanarao, K. A joint context-aware embedding for trip recommendations. In Proceedings of the 2019 IEEE 35th International Conference on Data Engineering, Macao, China, 8–11 April 2019; pp. 292–303.
11. Lim, K.H.; Chan, J.; Leckie, C.; Karunasekera, S. Personalized tour recommendation based on user interests and points of interest visit durations. In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina, 25–31 July 2015; pp. 1778–1784.
12. Yahi, A.; Chassang, A.; Raynaud, L.; Duthil, H.; Chau, D.H. Aurigo: An interactive tour planner for personalized itineraries. In Proceedings of the 20th International Conference on Intelligent User Interfaces, Atlanta, GA, USA, 29 March–1 April 2015; pp. 275–285.
13. Xiang, Z.; Gretzel, U. Role of social media in online travel information search. *Tour. Manag.* **2010**, *31*, 179–188. [CrossRef]
14. Kurashima, T.; Iwata, T.; Irie, G.; Fujimura, K. Travel route recommendation using geotags in photo sharing sites. In Proceedings of the 19th ACM International Conference on Information and Knowledge Management, Toronto, ON, Canada, 26–30 October 2010; pp. 579–588.
15. Ouyang, Y.; Liu, W.; Rong, W.; Xiong, Z. Autoencoder-based collaborative filtering. In Proceedings of the International Conference on Neural Information Processing, Montreal, QC, Canada, 6 June 2014; pp. 284–291.
16. Sedhain, S.; Menon, A.K.; Sanner, S.; Xie, L. Autorec: Autoencoders meet collaborative filtering. In Proceedings of the 24th International Conference on World Wide Web, New York, NY, USA, 18–22 May 2015; pp. 111–112.
17. Strub, F.; Gaudel, R.; Mary, J. Hybrid recommender system based on autoencoders. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, Boston, MA, USA, 15 September 2016; pp. 11–16.
18. Okazaki, S.; Andreu, L.; Campo, S. Knowledge sharing among tourists via social media: A comparison between Facebook and TripAdvisor. *Int. J. Tour. Res.* **2017**, *19*, 107–119. [CrossRef]
19. Chae, D.K.; Kang, J.S.; Lee, J.T.; Kim, S.W. CFGAN: A generic collaborative filtering framework based on generative adversarial networks. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Turin, Italy, 22–26 October 2018; pp. 38–40.
20. Robusto, C.C. The cosine-haversine formula. *Am. Math. Mon.* **1957**, *64*, 38–40. [CrossRef]
21. Beel, J.; Langer, S.; Genzmehr, M.; Gipp, B.; Breitinger, C.; Nürnberger, A. Research paper recommender system evaluation: A quantitative literature survey. In Proceedings of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation, Hong Kong, China, 12 October 2013; pp. 15–22.
22. Chae, D.K.; Kim, S.W.; Lee, J.T. Autoencoder-based personalized ranking framework unifying explicit and implicit feedback for accurate top-N recommendation. *Knowl.-Based. Syst.* **1957**, *64*, 38–40. [CrossRef]
23. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.