Booking.com RecSys RecTour 2024 Challenge

Amit Livne[1], Eran Fainman[2]

Abstract

The ACM RecSys RecTour 2024 Challenge focuses on ranking reviews, which is an important aspect that influences users' decision-making. The goal of this challenge is to match given accommodations and users to their respective review IDs. The concept is that when a new user interacts with the booking system, we can analyze the accommodation they are viewing along with available user features (e.g., couple, country, etc.). This enables us to display reviews in an order that considers the review content with respect to the user and accommodation characteristics. To do so, Booking.com provides a unique training dataset containing 1.6 million reviews based on real anonymized bookings. Thirty teams signed up for the challenge. The top-performing teams are invited to present short papers describing their solution approaches.

Keywords

Tourism recommenders, Ranking, Reviews

1. Problem Description

Our objective is to create a model that predicts the helpfulness of every review tailored to individual users. In essence, we aim to construct a personalized helpfulness function, denoted as $f(r_j|c_i)$, which evaluates the relevance of review r_j to user i given their context c_i. This function assigns a score indicating the degree to which review j is beneficial for user i. These scores enable us to rank reviews, ensuring that those with the highest f values are deemed most helpful within context c_i.

Using the number of helpful votes as the target signal inherits multiple issues. First, it introduces a presentation bias towards the previous review ranking algorithm (usually sorted by votes). Additionally, the signal of votes is sparse as most of the reviews are not presented and therefore not voted. Moreover, there might be a cold-start problem where new reviews don't have as many votes as older reviews which might be less relevant over time. Finally, in many cases, only the final number of votes is stored and therefore it's not feasible to use this signal for developing personalized review ranking models.

Thus, we introduce a more feasible and novel approach for modeling personalized helpfulness measure. We model the personalized helpfulness of a review as the likelihood that it is written by its reviewer given the reviewer's context. Notably, we define f such that given a user context c_i, it estimates the likelihood that review r_j was written by the user. Formally, we optimize f

Workshop on Recommenders in Tourism (RecTour 2024), October 18th, 2024, co-located with the 18th ACM Conference on Recommender Systems, Bari, Italy

amit.livne@booking.com (A. Livne); eran.fainman@booking.com (E. Fainman)

such that given that review r_i was written by a user with context c_i, it holds:

$$f(r_j|c_i) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (1)$$

In this challenge the task is to match given accommodations and users to their respective review IDs. Therefore, we provide three sets of data as follows:

Users – hold information regarding users and accommodation features. Review – hold information regarding reviews. Matches – a true label between given *user_id*, *accomodaiton_id* and *review_id* (only positive examples).

2. Dataset

The dataset we publish contains authentic user-generated reviews from 50,000 accommodations. This includes information on the user reservation, the review and the accommodation. The dataset consists of 2,031,914 anonymized reviews, along with guest and accommodation context. It is based on real data, and available via the following repo[1]. Table 1 describes the dataset fields. Further information regarding the creation of the dataset is presented in [1, 2].

3. Challenge Timeline

Key dates of the challenge are listed in Table 2.

4. Submission Guidelines

The submission file should include 12 columns: Accommodation ID, User ID, and the top 10 review IDs ranked according to your model's predictions on the test set.

An example of a ranked review for Accommodation ID 1 and User ID 1. The algorithm predicted the following ranking of reviews (158, 32, ... 97). Consequently, the submission file will display it as demonstrated by Table 3.

The top 10 teams will be invited to submit short papers (up to 4 pages + references). The papers will include the team and the authors names, an abstract, sections describing the method and results, a link to their code repository and a reference to the Booking.com challenge in the following format: Amit Livne, and Eran Fainman. 2024. Booking.com RecSys RecTour 2024 Challenge. http://www.bookingchallenge.com , In Workshop on Recommenders in Tourism (RecTour 2024), October 18th, 2024, co-located with the 18th ACM Conference on Recommender Systems, Bari, Italy.

Selected papers are expected to present their work in the workshop. The submitted papers will be evaluated based on their clarity, novelty, and results presentation.

Please contact rectour2024challenge@booking.com for any questions.

[1]https://huggingface.co/datasets/Booking-com/accommodation-reviews

Table 1
Dataset description fields

Field Name	Description
review_title	Review title
review_positive	”liked” section of the review - textual information
review_negative	”disliked” section of the review - textual information
review_score	Overall review score for the stay
review_helpful_votes	How many users marked the review as helpful
guest_type	There are 4 traveller types: Solo traveller (1 adult) / Couple (2 adults) / Group (>2 adults) / Family with children
guest_country	Anonymized guest country
room_nights	Number of nights booked
month	The month of the check-in date of the reservation
accommodation_id	Anonymized accommodation ID
accommodation_type	hotel/ apartment/ hostel
accommodation_score	average guest review score of the accommodation
accommodation_country	Country of the accommodation
location_is_beach	Is the accommodation located in a beach location
location_is_ski	Is the accommodation located in a ski location
location_is_city_center	Is the accommodation located in a city center location

5. Evaluation Criteria and Leaderboard

The goal of the challenge is to predict (and recommend) the review ID of each Accommodation and User IDs pair. The quality of the predictions is evaluated based on the MRR@10. The online results are available via the leaderboard[2]

[2]https://huggingface.co/spaces/Booking-com/rectour24-review-ranking-leaderboard-test

Table 2
Challenge key dates

When?	What?
June 20th, 2024	Registration and challenge start
August 10th, 2024	Release of validation dataset
August 27th, 2024	Release of test dataset
September 1st, 2024	Announcement on the winners
September 3rd, 2024	Paper submission deadline
September 20th, 2024	Paper notifications
September 27th, 2024	Camera-ready submissions due
October 19th, 2024	RecTour 2024 takes place

Table 3
Submission format

Accomodation ID	User ID	Review 1	Review 2	...	Review 10
1	1	158	32	..	97

6. Results

60 participants have signed up for the challenge. After two months of a contest, 10 of them applied a final submission. Top 4 performing teams are listed in Table 4.

Table 4
Top performing teams

	Team	**MRR@10**
1	ringo	0.1662
2	TMU-Rec	0.0775
3	BMS Hunters	0.0735
4	qtravel.ai	0.0735

The best performing team achieved MRR@10 of 0.1662. The teams have submitted short papers and code repositories with a detailed description of their solution methodology.

References

[1] R. Igebaria, E. Fainman, S. Mizrachi, M. Beladev, F. Wang, Enhancing travel decision-making: A contrastive learning approach for personalized review rankings in accommodations, arXiv preprint arXiv:2407.00787 (2024).

[2] F. Wang, M. Beladev, O. Kleinfeld, E. Frayerman, T. Shachar, E. Fainman, K. L. Assaraf, S. Mizrachi, B. Wang, Text2topic: Multi-label text classification system for efficient topic detection in user generated content with zero-shot capabilities, in: 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, 2023.