Booking.com WSDM WebTour 2021 Challenge

Dmitri Goldenberg
dima.goldenberg@booking.com
Booking.com, Tel Aviv, Israel

Kostia Kofman
kostia.kofman@booking.com
Booking.com, Tel Aviv, Israel

Pavel Levin
pavel.levin@booking.com
Booking.com, Amsterdam, Netherlands

Sarai Mizrachi
sarai.mizrachi@booking.com
Booking.com, Tel Aviv, Israel

Maayan Kafry
maayan.kafry@booking.com
Booking.com, Tel Aviv, Israel

Guy Nadav
guy.nadav@booking.com
Booking.com, Tel Aviv, Israel

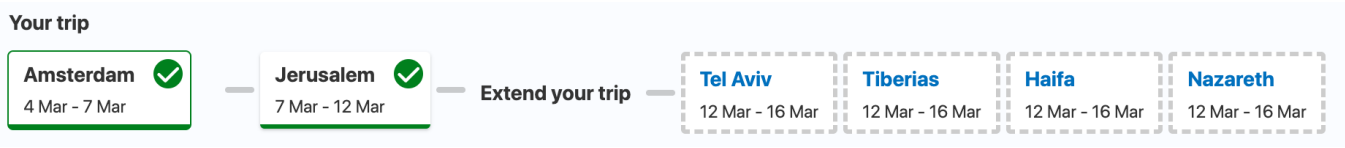

Figure 1: Multi-Destinations trip recommendation bar on Booking.com website

ABSTRACT

The ACM WSDM WebTour 2021 Challenge focuses on a multi-destinations trip planning problem, which is a popular scenario in the travel domain. The goal of the challenge is to make the best recommendation of an additional trip destination. To encourage research on this field, Booking.com provided a unique dataset based on millions of real anonymized bookings.

More than 800 participants have signed up for the contest. Best performing team achieved *Accuracy @ 4* of 0.5939, using a blend of Transformers, GRUs, and feed-forward multi-layer perceptron. Additional leading teams implemented advanced state-of-the-art solutions to tackle the problem.

CCS CONCEPTS

• **Information systems** → **Personalization**; **Recommender systems**.

KEYWORDS

Personalization, Travel, Recommender Systems, Dataset, Challenge

1 PROBLEM DESCRIPTION

Booking.com's mission is to make it easier for everyone to experience the world. By investing in the technology that helps take the friction out of travel, Booking.com seamlessly connects millions of travelers with memorable experiences, a range of transport options, and incredible places to stay.

Many of the travelers go on trips that include more than one destination. For instance, a user from the US could fly to Amsterdam for 5 nights, then spend 2 nights in Brussels, 3 in Paris, and 1 in Amsterdam again before heading back home. In this scenario, the users are offered personalized recommendations [1] for extending their trip immediately when they make their booking, as shown in figure 1.

The goal of this challenge is to use a dataset based on millions of real anonymized accommodation reservations to come up with a strategy for making the best *recommendation for their next destination* [2]. Booking.com releases this unique dataset to encourage the research on sequential recommendation problems [4]. The challenge is part of the WebTour 2021 ACM WSDM workshop [3] on web tourism that will be held at the 14th ACM international 2021 WSDM Conference.

Table 1: Dataset description

Column	Description
user_id	User ID
checkin	Reservation check-in date
checkout	Reservation check-out date
affiliate_id	An anonymized ID of affiliate channel where the booker came from (e.g. direct, 3rd party referrals, paid search engine, etc.)
device_class	desktop/mobile
booker_country	Country from which the reservation was made (anonymized)
hotel_country	Country of the hotel (anonymized)
city_id	city_id of the hotel's city (anonymized)
utrip_id	trip ID (a group of multi-destinations bookings within the same trip)

2 DATASET

The dataset consists of over a million anonymized hotel reservations, based on real data, is available on the challenge website[1] and described in table 1. Each reservation is a part of a customer's trip (identified by utrip_id), which includes at least four consecutive reservations. There are 0 or more days between check-out and check-in dates of two consecutive reservations. The evaluation dataset is constructed similarly. However, the city_id of the final reservation of each trip is concealed and requires a prediction.

[1] https://www.bookingchallenge.com/

Table 2: Challenge key dates

When?	What?
December 1st, 2020	Challenge starts
January 6th, 2021	Test set release - teams registration deadline
January 28th, 2021	Challenge closes - results submission deadline
February 4th, 2021	Announcement on the winners
February 18th, 2021	Paper submission deadline
February 25th, 2021	Paper notifications
March 4th, 2021	Camera ready submissions
March 12th, 2021	Workshop day

3 CHALLENGE TIMELINE

Key dates of the challenge are listed in table 2.

4 SUBMISSION GUIDELINES

The teams are expected to submit their top four cities predictions per each trip on the test set until January 28th 2021. The submission will be done in a csv file named *submission.csv* in the following format described in table 3:

Table 3: Submission format

utrip_id	city_id_1	city_id_2	city_id_3	city_id_4
1000031_1	8655	8652	4323	4332

utrip_id represents each unique trip in the test, and the rest of the columns represent the *city_id* of the top four predicted cities. The top 10 teams will be invited to submit short papers (up to 4 pages + references in ACM sigconf format[2]). The papers will include the team and the authors names, an abstract, a text describing the method and the achieved score, and a link to their code repository and refer to the Booking.com challenge in the following format:

```
Dmitri Goldenberg, Kostia Kofman, Pavel Levin, Sarai
Mizrachi, Maayan Kafry, and Guy Nadav. 2021. Booking.com
WebTour 2021 Challenge. http://www.bookingchallenge.com
, In ACM WSDM Workshop on Web Tourism (WSDM WebTour'21),
March 12th, 2021, Jerusalem, Israel.
```

Selected papers are expected to present their work in the workshop (in a virtual format). The submitted papers will be evaluated based on their clarity, novelty, and results presentation. Please contact *wsdmchallenge@booking.com* for any questions.

5 EVALUATION CRITERIA

The goal of the challenge is to predict (and recommend) the final city (city_id) of each trip (utrip_id). The quality of the predictions is evaluated based on the top four recommended cities for each trip by using *Top-4 Accuracy metric* (4 representing the four suggestion slots at Booking.com website). When the true city is one of the top 4 suggestions (regardless of the order), it is considered correct.

[2]https://www.acm.org/publications/proceedings-template

Table 4: Top 10 performing teams

	Team	Accuracy @ 4
1	**NVIDIA RAPIDS.AI**	**0.5939**
2	**Synerise AI**	**0.5780**
3	**TEAM DASOU**	**0.5741**
4	mbaigorria	0.5566
5	aprec	0.5557
6	hakubishin3 & u++ & yu-y4	0.5399
7	testing	0.5332
8	Alexander Makeev	0.5310
9	YiNet	0.5112
10	Marlesson - MARS-Gym	0.4958

6 PRIZES

To encourage research contributions, the top three performing teams will receive Booking.com Travel Credits. The best paper team will receive an additional prize. Paper submission and virtual participation at the workshop are mandatory in order to be eligible for a prize.

7 RESULTS

820 participants have signed up for the challenge. After two months of a contest, 97 of them performed a final submission, grouped in 40 competing teams. Top 10 performing teams are listed in table 4. The best performing team achieved *Accuracy @ 4* of 0.5939, implementing a blend of 3 different neural network architectures, using Transformers, GRUs, and feed-forward MLP. Other solutions relied on Efficient Manifold Density Estimator, LSTM networks, Attention mechanisms, Lambdarank, and further state-of-the-art methods. The teams have submitted short papers and code repositories with a detailed description of their solution methodology.

REFERENCES

[1] Dmitri Goldenberg, Kostia Kofman, Javier Albert, Sarai Mizrachi, Adam Horowitz, and Irene Teinemaa. 2021. Personalization in Practice: Methods and Applications. In *Proceedings of the 14th International Conference on Web Search and Data Mining.*

[2] Julia Kiseleva, Melanie JI Mueller, Lucas Bernardi, Chad Davis, Ivan Kovacek, Mats Stafseng Einarsen, Jaap Kamps, Alexander Tuzhilin, and Djoerd Hiemstra. 2015. Where to go on your next trip? Optimizing travel destinations based on user preferences. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval.* 1097–1100.

[3] Tsvi Kuflik, Catalin Mihai Barbu, Amra Delić, Dmitri Goldenberg, Julia Neidhardt, Ludocik Coba, and Markus Zanker. 2021. WebTour 2021 Workshop on Web and Tourism. In *Proceedings of the 14th International Conference on Web Search and Data Mining.*

[4] Sarai Mizrachi and Pavel Levin. 2019. Combining Context Features in Sequence-Aware Recommender Systems.. In *RecSys (Late-Breaking Results).* 11–15.