A Re-rank Algorithm for Online Hotel Search

Emanuele Cavenaghi[1]([✉]), Lorenzo Camaione[4], Paolo Minasi[4],
Gabriele Sottocornola[1], Fabio Stella[3], and Markus Zanker[1,2]

[1] Free University of Bozen-Bolzano, Bolzano, Italy
{ecavenaghi,gabriele.sottocornola,Markus.Zanker}@unibz.it
[2] University of Klagenfurt, Klagenfurt am Wörthersee, Austria
[3] University of Milano-Bicocca, Milano, Italy
fabio.stella@unimib.it
[4] Bravonext SA t/a lastminute.com, Chiasso, Switzerland
{lorenzo.camaione,paolo.minasi}@lastminute.com

Abstract. Recommender Systems were created to support users in situations of information overload. However, users are consciously or unconsciously influenced by several factors in their decision-making. We analysed a historical dataset from a meta-search booking platform with the aim of exploring how these factors influence user choices in the context of online hotel search and booking. Specifically, we focused our study on the influence of (i) ranking position, (ii) number of reviews, (iii) average ratings and (iv) price when analysing users' click behaviour. Our results confirmed conventional wisdom that position and price were the "two elephants in the room" heavily influencing user decision-making. Thus, they need to be taken into account when, for instance, trying to learn user preferences from clickstream data. Using the results coming from this analysis, we performed an online A/B test on this meta-search booking platform comparing the current policy with a price-based re-rank policy. Our online experiments suggested that, although in offline experiments items with lower prices tend to have a higher Click-Through Rate, in an online context a price-based re-rank was only capable to improve the Click-Through Rate metric for the first positions of the recommended lists.

Keywords: Recommender systems · Tourism · Meta-search booking platform · Online hotel search · Data analysis · Learning to rank

1 Introduction

Recommender Systems (RSs) [7] are algorithms developed for helping users to find items of interest. The massive volume of information available on the web leads to the problem of information overload and, thus, increases the need for delivering effective and timely recommendations. The main idea behind these methods is to estimate users' interests from their past interactions in order to recommend them new unseen items matching their preferences.

As recently surveyed in [18], RSs are extensively applied in the E-Tourism domain to recommend destinations/travel packages [23], points of interest [22] or restaurants [24]. When recommending accommodations it is fundamental to exploit both contextual features (such as season and place) as well as users' preferences. Over the last years, many RSs were developed in order to recommend hotels in the context of online booking. Some works applied traditional RSs techniques such as *Collaborative Filtering* [8,15] or *Content-Based approaches* [17,20]. Instead, other works proposed domain-specific approaches. For instance, [10] use textual reviews as the main source of information to make recommendations, [19] build specific topic models from textual reviews or [11] design an app where users can search and browse hotel reviews. Finally, another approach to recommend hotels is the *Learning to Rank* approach (see [12]) in order to automatically construct a ranking model using training data, such that the model can sort items according to their relevance, preference, or importance (e.g. [14]).

One work studying the influence of different factors on the user decision-making is the one in [1]. Here, the authors examined the impact of three main factors on converting searches into customers: impact of rated reviews, recommendations and search listings. They showed that a high rank position in the list, room price and hotel size had a significant effect on conversion rates. Instead, the rating of the location had a significant positive impact on conversion rates while service rating and star rating did not show a significant effect. These findings suggest that ratings are not always reliable as also found by [16]. They examined whether encountering reviews shared on social media containing disparity in the cost of same accommodation could cause regret and alter the intention to revisit from a retrospective point of view.

The first contribution of our work is to study the influence of several observable variables related to online hotel search in the context of a meta search engine, i.e. a platform that aggregates data from multiple *Online Travel Agencies* (OTAs)[1]. Specifically, we focus our attention on the following features of the properties[2]: (i) the rank position of accommodations in the recommendation list displayed to the user, (ii) the price, (iii) the average rating and (iv) the number of reviews. Based on the analysis of the impact of the aforementioned properties on users' click behavior, we tested a simple re-rank algorithm and report results from an online A/B test. Therefore, the research questions are as follows:

– **R.Q. 1:** Do the variables (rank position, price, rating and number of reviews) influence user decision-making?
– **R.Q. 2:** Does a price-based re-rank of recommended items lead to a higher Click-Through Rate (CTR)?
– **R.Q. 3:** Does the OTA associated with each offered property influence the CTR?

To answer the first research question, we analysed an historical dataset collected on a *meta-search booking platform* that aggregates offered properties from

---

[1] The OTA is an external party which facilitates the booking a property.
[2] With the term property we refer to any type of accommodation like hotels, apartment houses, etc.

different OTAs. Instead, to answer the second and the third research questions, we ran an A/B test on the meta-search platform to compare the existing recommendation strategy of the meta-search platform with a local price-based re-rank algorithm. In both, the historical dataset and the A/B test, the company had no information about the anonymous users and their history of previous interactions with the site. Moreover, there was no explicit feedback (e.g., user ratings specific to properties), but we had to rely solely on implicit feedback [6], in our case user clicks.

2   Data Analysis

In this section, we report the results of our data analysis to provide an answer to the first research question: "Do the variables (rank position, price, rating and number of reviews) influence user decision-making?". The dataset has been collected on a meta-search booking platform, where the roughly $130,000$ recommended lists and the associated user actions (click-throughs) of different anonymous user sessions were recorded, each consisting of 25 properties/recommended accommodations displayed on the same page.

The data has been collected on searches about 14 Italian cities on a meta-search booking platform in the period between $11/2021$ and $4/2022$. Moreover, in searches where users applied a filter criterion we could not unambiguously map the clicks to the specific search and therefore we removed all searches with any filter applied. In every recommended list, each property can be associated with a different OTA. As a result, each property can be presented with different OTAs and in each list multiple OTAs are present. Given the small number of users that look beyond the first page, we restricted the analysis to the first page with 25 ranked properties.

Figure 1 reports the CTR computed for each of the 25 positions in the recommended list. In this and all the following figures, the shaded area around the line represents the 99% confidence interval computed by the *Seaborn*[3] python package using a bootstrapping technique. For the remainder of the paper, we will refer to the CTR distribution shown in Fig. 1 as the *a-priori CTR*. Instead, Figs. 2, 3 and 4, report the CTR distribution for values of price, rating and number of reviews that are higher than the 80% quantile and lower than the 20% quantile. For example, in Fig. 2, to obtain the CTR distribution for values higher than the 80% quantile, we took only the properties with a price higher than the 80% quantile into account within each recommended list, i.e. we removed all the properties with a price lower than the 80% quantile in each recommended list from the dataset. We chose this method to investigate the cheapest/best rated/most reviewed and most expensive/worst rated/least reviewed. We initially focused on the 10% and 90% quantiles, but due to the low number of samples we switched to 20% and 80% quantile in order to obtain statistically significant results.
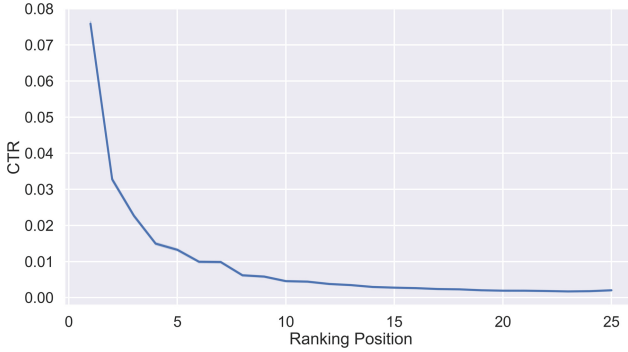
---

[3] https://seaborn.pydata.org/.

Fig. 1. CTR for each rank position in the recommendation list. The y-axis reports the CTR achieved in each rank position. Due to the large number of observations, the confidence interval is very small and difficult to see in the image.

Influence of Rank Position. It is well documented in the literature that the rank position affects user choice [4, 9], and our analysis confirmed this assertion. From Fig. 1 it was clear that the rank position had a significant effect on CTR. In fact, the CTR for rank position k was always higher than the CTR for subsequent rank positions. For example, in our dataset, the CTR for the first rank position was around 5.75%, while the CTR for the last rank position was lower than 0.5%. This means that the probability of clicking the first rank position was about 12 times higher than clicking on rank position 25. However, while for the first 10 rank positions this was evident, the difference in CTR for the rank positions between 10 and 25 was close to zero. The answer to the research question: "To which extent does the rank position influence user decision-making?" is yes: the rank position clearly influenced user's choice.

Influence of Price. Figure 2 reports the results for price when considering only properties having a price above the 80% quantile and below the 20% quantile. For each rank position, the CTR was higher than the a-priori CTR if we considered properties with a price lower than the 20% quantile. This clearly means that lower prices positively influenced the users' propensity of clicking on a property and the opposite happened if we considered properties with a price higher than the 80% quantile: for higher prices the CTR was lower. The answer to the question: "To which extent does the price influence user decision-making?" is clearly yes. Price influenced the user decision-making both positively and negatively as also stated in [13, 21].

Influence of Rating. Figure 3 reports the analysis for properties with a rating above the 80% quantile and below the 20% quantile. While the influence of price was statistically significant for every rank position, for rating this difference was only significant for properties with a rating below the 20% quantile and only for the first seven rank positions. We expected that a higher rating would lead
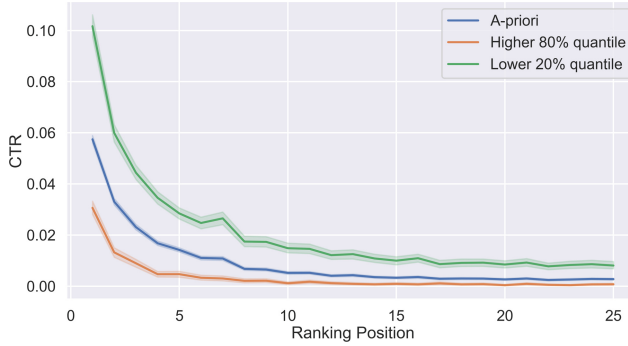
Fig. 2. CTR for each rank position in the recommendation list taking into account property's price.

to a higher CTR, while a lower rating would lead to a lower CTR. Instead, we observed the exact opposite phenomenon. Thus, the answer to the research question: "To which extent does the rating influence user decision-making?" is not so clear. While for rank position and price the effect was always relevant, for the rating it was more difficult to give a unique answer: the rating seemed to have a very weak influence or no influence at all. These results are in contrast with [5] and [2] that found that users were influenced by rating. However, these works ran user studies considering only rating and number of reviews, while for us price seemed to have a stronger influence that probably confounded the effect of rating.
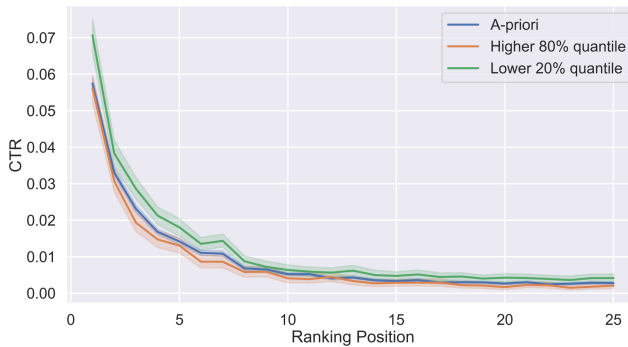

Fig. 3. CTR for each rank position in the recommendation list taking into account property's rating.

Influence of Number of Reviews. Finally, we report in Fig. 4 the analysis for properties with a number of reviews above the 80% quantile and below the 20% quantile. In these cases, it was clear that the number of reviews did not

influence the user's choice at all. These results are in contrast to other studies in the literature, such as [2], where users tend to trust the rating of the items only if they are made up on a sufficient number of reviews. To verify the existence of a sufficient number of reviews above which users trust the rating, we compared the distribution of the CTR for properties with above and below 35 reviews. We found that there was a positive influence for properties above the threshold confirming our claims. Given our data, the last research question: "To which extent does the number of reviews influence user decision-making?" has to be answered as follows: the number of reviews did not influence user decision-making when the number of the reviews exceeded a specific threshold.
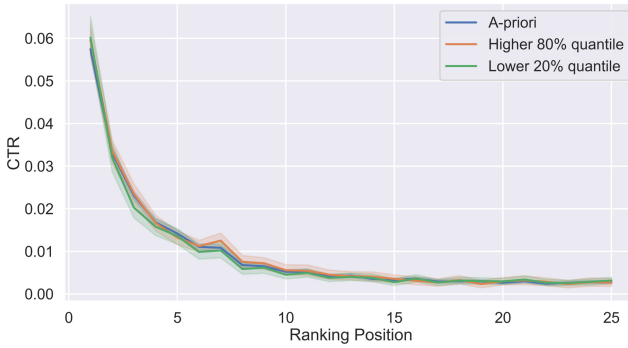

Fig. 4. CTR for each rank position in the recommendation list taking into account the number of reviews of each property.

3   Re-rank Algorithm

To answer the second research question, we implemented a simple and efficient algorithm to re-rank the top-25 list of offered properties as recommended by the current algorithm. Since the developed algorithm only re-ranks the top-25 items, it is ensured that all properties presented to users are of comparable quality with the baseline. To re-rank the properties, we computed a score and reordered the properties from highest to lowest score. The score, reported in Eq. 1, is composed by two logistic functions with two means:

$$y = \alpha \cdot \frac{1}{1 + e^{kx_i - \mu_1}} + \beta \cdot \frac{1}{1 + e^{kx_i - \mu_2}} \tag{1}$$

where $\alpha, \beta \in [0, 1]$ manage the weight of the two functions while $\alpha + \beta = 1$, y represents the score, x_i the price of the property i, and k controls the speed by which the function approaches the limits (i.e., 0 and 1). Finally, the two means, μ_1 and μ_2, represent respectively the mean price for the type of accommodation[4]

---

[4] With *type of accommodation* we refer to the different type of properties, e.g., apartment, guest house, hotels with 3 stars, etc.

of the property i within the recommendation list and the median price of the properties within the recommended list (regardless of the type of accommodation). For μ_2, we used the median instead of the mean to reduce the impact of outlier prices, for instance, the price of 5-stars hotels. μ_1 allows us to account in a simple way the quality-price ratio, because a user may prefer to pay more for higher quality accommodations. While μ_2 controls for the absolute price of the properties, because, as showed in Fig. 2, users tend to click on properties associated with a lower price. In the following experiments, we use $\alpha = \beta = 0.5$. We select these values using the results from offline experiments on the dataset described in Sect. 2, because running multiples online experiments with different values of the two hyper-parameters was not possible.

4 A/B Test Results

The A/B test was conducted on the company's website for 20 days (between June and July 2022), and, in the end, nearly 1 million searches were conducted by users worldwide.
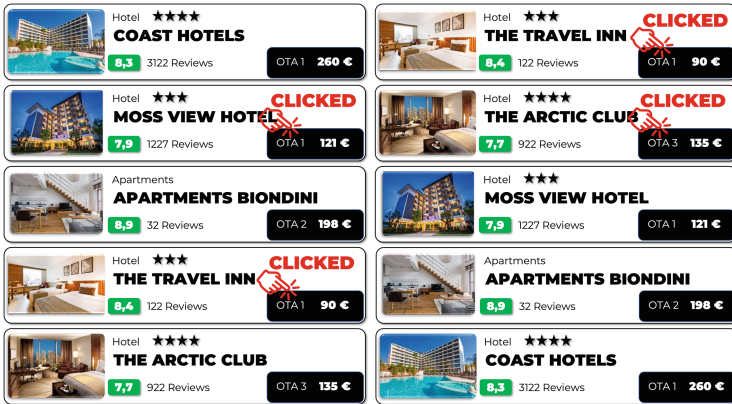

Fig. 5. Example of the A/B test. On the left is the list of properties recommended by the Baseline policy, and on the right is the list reordered by the Re-rank policy. The achieved NDCG are 0.651 (Baseline policy) and 1.0 (Re-rank policy) even if the number of clicks are the same.

We compared the Baseline policy used by the company, a linear regression that takes several factors into account and where weights are manually chosen by experts, with the Re-rank policy described in Sect. 3. Figure 5 reports an example of the A/B test on a 5 items list while the results, in terms of CTR for each rank position, are reported in Fig. 6.

Figure 6 clearly depicts that, for the first position, the CTR achieved by the Re-rank policy was statistically significantly higher (more than 2%) than the Baseline policy. Instead, for all rank positions after the third, the Baseline

policy achieved a slightly higher CTR, even if the difference was less than 0.5% and close to zero for bottom positions. The increase in the first position was expected, and the results confirmed our hypotheses. However, we also expected an improvement for more top ranked positions while from the third rank position we observed a decrease.

To further analyse the user click behaviour, given that we can not disclose the results in terms of conversion rates, we computed the CTR for search (SCTR). The SCTR is defined as the ratio between the number of clicked searches and the total number of searches: $\frac{\text{\# of clicked searches}}{\text{\# of searches}}$. A search is clicked if at least one of the recommended item received a click. The Re-rank policy achieved a SCTR of 23.48%, while the Baseline policy achieved a slightly higher SCTR of 24.16%. The difference between the two policies was very small and showed that the increase in CTR in the first position for the Re-rank policy was compensated by the decrease for all the other positions. Finally, we also computed the *Normalized Discounted Cumulative Gain (NDCG)* on the clicks in order to assess if the Re-rank policy improved the ranking quality w.r.t. the Baseline policy. The Baseline policy achieved a NDCG of 0.121 while the Re-rank policy improved the NDCG to 0.136. This further highlighted that the Re-rank policy was more effective in recommending more relevant properties in the top positions.
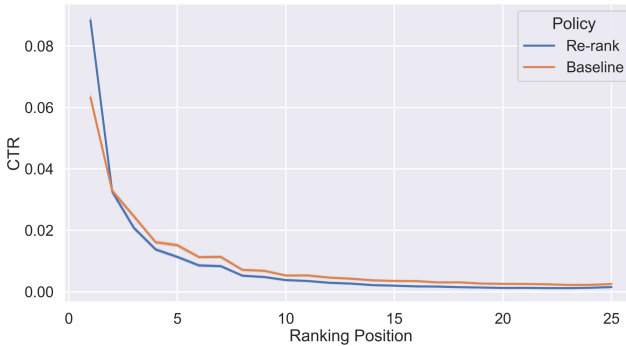

Fig. 6. Results from the A/B test in terms of CTR for each rank position. Due to the large number of observations, the confidence interval is very small and difficult to see in the image.

Given the results from the A/B test, the answer to the second research question, "Is a price-based re-rank of these offered properties sufficient to achieve a higher CTR?", is clearly yes if the main goal is to improve the CTR in the top position of the list. Despite the data analysis results, which showed that lower prices were a key factor to improve the CTR, a policy that re-rank items by price was only sufficient to improve the CTR w.r.t. the Baseline policy in the first position of the recommended list. However, since users usually pay more attention to the item in first position, this can be considered a good result even if the SCTR slightly decreased with the Re-rank policy.

5   Influence of the OTA

To further study the differences in CTR and SCTR metrics between the two policies, we analysed additional variables with potential influence on user decision-making. While the variables analysed before, i.e. average rating, number of reviews and price, behave similarly as before, the OTA presented with each property emerged as one of the key factors in the A/B test. Furthermore, this variable is very important to the company's business, so we focused our attention on it.
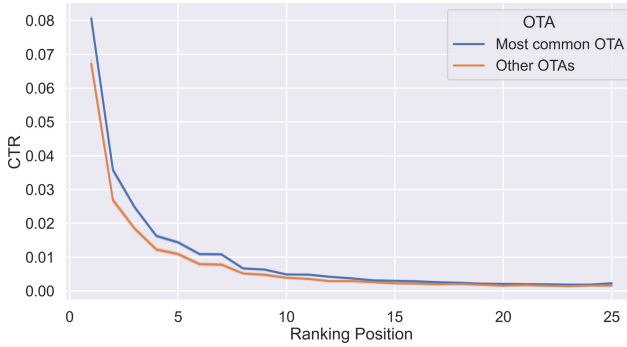

Fig. 7. CTR comparison for each rank position between the *most common OTA* and the *other OTAs*. Due to the large number of observations, the confidence interval is very small and difficult to see in the image.

Figure 7 depicts the CTR at each rank position for the *most common OTA* and for all the *other OTAs*. Since we can not disclose the names of the OTAs, we only distinguish between the *most common OTA* and the *other OTAs*. The *most common OTA* always had a CTR that was significantly higher than the CTR of the *other OTAs*, at least for the first 15 positions, which means that users preferred this OTA to the others. One reason for this preference could be that the *most common OTA* might be more trusted by users.

This difference in CTR between OTA groups, joined with the number of recommendations for each OTA group at each rank position, reported in Fig. 8, could explain the difference in SCTR identified between the two policies. From Fig. 8a, we can see that the Re-rank policy recommended properties with the *most common OTA* less frequently at top-ranked positions while favouring more frequently *other OTAs*, Fig. 8b.

Thus, by favouring lower price offers the Re-rank policy pushed less well-known OTAs to top-ranked positions and exposed them to higher levels of users' attention. Their lower likelihood of being clicked, however, seems to have neutralized the positive price effect and resulted in an overall decrease in the SCTR.
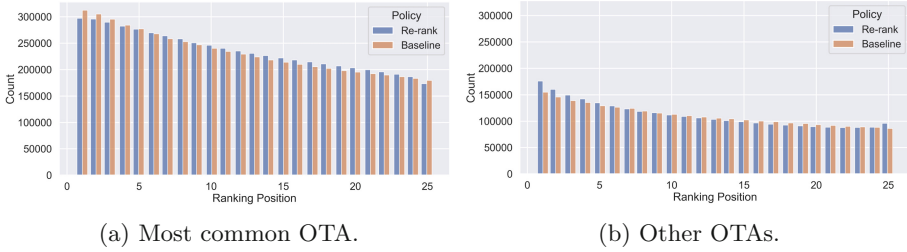
(a) Most common OTA.   (b) Other OTAs.

Fig. 8. Count of recommendations for the *most common OTA* and for the *other OTAs*.

The answer to the third research question, "Does the OTA associated with each offered property influence the CTR?", is yes. Although price and rank positions were identified as the most important features that influenced users' decision-making, there were also other factors, in our case the OTA, that could impact users' decision and thus overall performance metrics of a ranking policy. At the end, in our case, a price-based re-rank algorithm that also keep the balance for the OTA feature would probably have improved the baseline, whereas considering only the price was sufficient to achieve a marginal improvement.

6   Conclusions

In this paper, we studied how different variables affect the user click behaviour in online hotel search. Specifically, we took into account the following variables: (i) rank position, (ii) price, (iii) rating and (iv) number of reviews, and measured their influence by observing changes in the CTR distribution. We started by analysing a historical dataset collected in a *meta-search booking platform* in which as expected the rank position and the price heavily influenced the user's choice: a property in the top positions had a greater probability of receiving a click than a property in the last positions and a high price generally discouraged users.

On the other hand, differently from the previous literature, we found that the average rating had a weak influence on the user's choice. Probably, this influence was heavily confounded by the influence of the price because, as stated by [3], average ratings are influenced by price: products tend to have a higher rating when they have a higher price. Even regarding the number of reviews, we discovered that this variable did not have a significant effect on user's choice. This was because above a threshold, about 35 reviews in our case, the user trusted the average rating, and a larger number of reviews did not change the user's perception.

To further verify the influence of price, we ran an online A/B test on the company's website to compare a Baseline policy with a price-based re-rank policy that shuffles the top-25 offered properties in recommendation lists. The results showed that the re-rank policy improved the CTR for the first rank position,

confirming that price was a key factor that influenced users click behaviour. Furthermore, we observed that also the OTA associated with a property, influenced the user decision. For example, in our case the *most common OTA* achieved a higher CTR for every rank position compared to the *other OTAs* and seems to be favoured more by users.

This work consequently highlights the many influence factors and biases on users decision-making in online travel search that are disregarded in most offline datasets by presenting the outcome of a price-based re-rank strategy.

References

1. Cezar, A., Ögüt, H.: Analyzing conversion rates in online hotel booking: the role of customer reviews, recommendations and rank order in search listings. Int. J. Contemp. Hosp. Manag. **28**(2), 286–304 (2016)
2. Coba, L., Rook, L., Zanker, M.: Choosing between hotels: impact of bimodal rating summary statistics and maximizing behavioral tendency. Inf. Technol. Tourism **22**(1), 167–186 (2020)
3. De Langhe, B., Fernbach, P.M., Lichtenstein, D.R.: Navigating by the stars: investigating the actual and perceived validity of online user ratings. J. Consum. Res. **42**(6), 817–833 (2016)
4. Feigenbaum, E.A., Simon, H.A.: A theory of the serial position effect. Br. J. Psychol. **53**(3), 307–320 (1962)
5. Gavilan, D., Avello, M., Martinez-Navarro, G.: The influence of online ratings and reviews on hotel booking consideration. Tour. Manage. **66**, 53–61 (2018)
6. Jannach, D., Lerche, L., Zanker, M.: Recommending based on implicit feedback. In: Brusilovsky, P., He, D. (eds.) Social Information Access. LNCS, vol. 10100, pp. 510–569. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90092-6_14
7. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: Recommender Systems: an Introduction. Cambridge University Press, Cambridge (2010)
8. Jannach, D., Zanker, M., Fuchs, M.: Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. Inf. Technol. Tourism **14**(2), 119–149 (2014). https://doi.org/10.1007/s40558-014-0010-z
9. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F., Gay, G.: Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. ACM Trans. Inf. Syst. (TOIS) **25**(2), 7–es (2007)
10. Levi, A., Mokryn, O., Diot, C., Taft, N.: Finding a needle in a haystack of reviews: cold start context-based hotel recommender system. In: Proceedings of the Sixth ACM Conference on Recommender Systems, pp. 115–122 (2012)
11. Lin, K.P., Lai, C.Y., Chen, P.C., Hwang, S.Y.: Personalized hotel recommendation using text mining and mobile browsing tracking. In: 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 191–196. IEEE (2015)
12. Liu, T.Y.: Learning to rank for information retrieval. Found. Trends Inf. Retr. **3**(3), 225–331 (2009)
13. Lockyer, T.: The perceived importance of price as one hotel selection dimension. Tour. Manage. **26**(4), 529–537 (2005)
14. Ludewig, M., Jannach, D.: Learning to rank hotels for search and recommendation from session-based interaction logs and meta data. In: Proceedings of the Workshop on ACM Recommender Systems Challenge. RecSys Challenge 2019. Association for Computing Machinery, New York (2019)

15. Nilashi, M., bin Ibrahim, O., Ithnin, N., Sarmin, N.H.: A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (EM) and PCA–ANFIS. Electron. Commer. Res. Appl. **14**(6), 542–562 (2015)
16. Nong, S.Z., Fong, L.H.N.: "Better Not Let Me Know": the mediating role of regret on the relation between social comparison discrepancy in online hotel review and revisit intention. In: Stienmetz, J.L., Ferrer-Rosell, B., Massimo, D. (eds.) ENTER 2022, pp. 87–96. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-94751-4_9
17. O'Mahony, M.P., Smyth, B.: Learning to recommend helpful hotel reviews. In: Proceedings of the Third ACM Conference on Recommender Systems, pp. 305–308 (2009)
18. Ricci, F.: Recommender systems in tourism. In: Handbook of e-Tourism, pp. 1–18 (2020)
19. Rossetti, M., Stella, F., Zanker, M.: Analyzing user reviews in tourism with topic models. Inf. Technol. Tourism **16**(1), 5–21 (2015). https://doi.org/10.1007/s40558-015-0035-y
20. Srisuan, J., Hanskunatai, A.: The ensemble of Naïve Bayes classifiers for hotel searching. In: 2014 International Computer Science and Engineering Conference (ICSEC), pp. 168–173. IEEE (2014)
21. Stávková, J., Stejskal, L., Toufarová, Z., et al.: Factors influencing consumer behaviour. ZEMEDELSKA EKONOMIKA-PRAHA **54**(6), 276 (2008)
22. Yang, W.S., Hwang, S.Y.: iTravel: a recommender system in mobile peer-to-peer environment. J. Syst. Softw. **86**(1), 12–20 (2013)
23. Yu, Z., Xu, H., Yang, Z., Guo, B.: Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints. IEEE Trans. Hum. Mach. Syst. **46**(1), 151–158 (2015)
24. Zeng, J., Li, F., Liu, H., Wen, J., Hirokawa, S.: A restaurant recommender system based on user preference and location in mobile environment. In: 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), pp. 55–60. IEEE (2016)