Analysis of Relevant Factors in Online Hotel Recommendation Through Causal Models

Emanuele Cavenaghi[1,*], Alessio Zanga[2,3], Alessandro Rimoldi[4], Paolo Minasi[4], Fabio Stella[2] and Markus Zanker[1,5]

[1]*Free University of Bozen-Bolzano, Bolzano, Italy*
[2]*University of Milano-Bicocca, Milano, Italy*
[3]*F. Hoffmann - La Roche Ltd, Basel, Switzerland*
[4]*Bravonext SA t/a lastminute.com, Chiasso, Switzerland*
[5]*University of Klagenfurt, Austria*

Abstract
Recommender systems were created to support users in situations of information overload. However, users are consciously or unconsciously influenced by many factors when making decisions, and the recommender must account for these to be effective. In this work, we use a causal graph to investigate the influence of different factors on the user's decision to click or not on the recommended accommodations. To learn the causal graph, we combine data provided by a meta-search booking platform for online hotel searches with prior knowledge made available by domain experts. The analysis confirms that the learnt causal model correctly models the well-known effect of the ranking position and price on user decision-making. Furthermore, we discover some interactions between the considered factors. For example, the country of the user market influences the user's decisions for different values of the price.

Keywords
Recommender Systems, Causal Networks, Tourism, Meta-search Booking Platform, Online Hotel Search

1. Introduction

Recommender Systems (RSs) [1, 2] have assumed a crucial role in several online services, such as e-commerce, entertainment and e-tourism. The massive volume of information on the web leads to the problem of information overload, which increases the need for delivering effective and timely recommendations. Indeed, large companies like Google, Facebook, Amazon, and Netflix have recognised that RSs are an essential part of their business [3, 4].

In the E-Tourism domain, RSs are extensively applied to recommend destinations/travel packages [5], points of interest [6] or restaurants [7]. Over the last years, many RSs have

RecSys Workshop on Recommenders in Tourism (RecTour 2023), September 19th, 2023, co-located with the 17th ACM Conference on Recommender Systems, Singapore
*Corresponding author.
ecavenaghi@unibz.it (E. Cavenaghi); alessio.zanga@unimib.it (A. Zanga); alessandro.rimoldi@lastminute.com (A. Rimoldi); paolo.minasi@lastminute.com (P. Minasi); fabio.stella@unimib.it (F. Stella); Markus.Zanker@unibz.it (M. Zanker)
0000-0002-0235-0421 (E. Cavenaghi); 0000-0003-4423-2121 (A. Zanga); 0000-0002-1394-0507 (F. Stella); 0000-0002-4805-5516 (M. Zanker)

been developed to recommend hotels in the context of online booking. Some works applied traditional RSs techniques such as *Collaborative Filtering* [8, 9] or *Content-Based approaches* [10, 11], while others proposed domain-specific approaches. For instance, [12] uses textual reviews as the main source of information to make recommendations, [13] builds specific topic models from textual reviews and [14] uses the *Learning to Rank* approach that is based on a ranking model to sort items according to their relevance or preference [15].

In hotel recommendations, it is fundamental to exploit contextual features (e.g., season and place) as well as users' features and preferences (e.g., age and nationality). Therefore, many works in the E-Tourism context studied the influence of several factors on user decision-making. For instance, [16] analyses the influence of several factors: ranking position, price, average rating and number of reviews. In the same spirit, [17] proposes a theory of the serial position effect and [18] studies the effect of the ranking position. Instead, [19, 20] study the influence of the item's price and conclude that lower prices positively influence user decision-making while higher prices have a negative influence. In [21], the authors performed a controlled user study to assess the effect of the items' average rating and the number of reviews on user decisions and concluded that both factors influence users.

However, these works analyse the influence of a single factor on user decision-making or, as in [21], analyse the influence in a controlled setting. Instead, in this work, we study the influence of different factors on users' decisions by analysing a historical dataset collected in the context of Online Hotel Search. Specifically, our dataset was collected on a *meta-search booking platform* that compares the prices of offered properties[1] from different *Online Travel Agencies* (OTAs)[2]. We propose an analysis using the causal framework [22, 23] to describe the problem and to analyse which factors of user, context and items influence the users' decision to click or not on the recommended properties. Specifically, we learned a *Causal Graph* (CG) and fitted a *Causal Network* (CN) to use it as an estimator. Causal approaches have also been applied to the RS problem in the last few years. However, to the best of our knowledge, the only work that faces the problem of learning a CG in RSs is [24]. Therefore, the main contributions of this paper are as follows:

- We describe the prior knowledge derived from the company's experts and previous works in the literature in terms of "tiers",
- We study which factors influence user decisions by learning a CG which combines observational data with prior knowledge given by domain experts,
- We analyse the relations between the factors and the *Click-Through Rate* CTR using the fitted CN model.

This paper is organized as follows, in Section 2, we report some important insights on the analysed context and the observational dataset. Then, in Section 3, we describe the process of learning the CG by combining observational data with expert knowledge and, in Section 4, we report some insights emerging from the performed data analysis exploiting the learnt CN.

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

2. Problem Statement

Firstly, it is important to underline the importance of analysing and investigating the context of interest in order to build up a coherent description of the problem using the causal framework. Therefore, in this section, we briefly describe the problem and the involved factors. We studied the problem of online hotel recommendation in a meta-search booking platform that compares offers from different OTAs for the same property. On this platform, users are not tracked, and thus, we have no information about previous interactions between a user and the platform. Moreover, the unique feedback we can exploit is the user's click on a recommended property. Therefore, we analysed the collected dataset by focusing on the *Click-Through Rate* (CTR), i.e. $\frac{\#\, of\, clicked\, properties}{\#\, of\, recommended\, properties}$.

The data has been collected on worldwide searches in the period between 11/2021 and 10/2022, where the roughly 8,200,000 recommended lists and the associated user actions (click-throughs) of different anonymous user sessions were recorded. Given the small number of users that look beyond the first page, we restricted the analysis to the first page, consisting of up to 25 ranked/recommended properties. Moreover, in searches where users applied a filter criterion, we could not unambiguously map the clicks to the specific search, and thus, we removed all searches with any filter applied.

In our analysis, we took into account factors of users, context and items to understand the relations between them and the user clicking on a property possessing certain features. Specifically, the included factors, partitioned into users' features, context's features and items' features, are reported in Table 1.

Table 1
Factors partitioned into users' features, context's features and items' features.

	Name	Description	Values
Users	POS	User's market country	{IT, FR, PT, US, ...}
	User cluster	User type	{single, couple, family}
	Device	Device used for the search	{mobile, desktop, tablet}
Context	People count	Number of People	[1,20]
	Rooms count	Number of Rooms	[1,4]
	Check-in month	Month of check-in	[1,12]
	Booking Window	Days between research and stay	[1, ∞]
	Length of Stay	Number of nights of stay	[1, ∞]
	State	Destination country	{Italy, France, Spain, ...}
	Properties in city	Number of properties in destination city	[1, ∞]
Items	OTA	Online Travel Agency	-
	TOA	Type Of Accommodation	{Hotel, Houses, Apartments, ...}
	Image count	Number of images in the gallery	[0, 7]
	Review count	Number of reviews of the property	[0, ∞]
	Rating	Average rating of the property	[0, 100]
	Price	Price of the property	[0, ∞]

3. Causal Discovery

A *probabilistic graphical model* (PGM) [25] is a tuple $(\mathcal{G}, \mathbf{X})$, where $\mathcal{G} = (\mathbf{V}, \mathbf{E})$ is a graph and $\mathbf{X}$ is a vector of random variables s.t. each vertex $V_i \in \mathbf{V}$ is associated to a random variable $X_i \in \mathbf{X}$. The graph $\mathcal{G}$ is said to be a *structure* over the associated *joint probability distribution* $P(\mathbf{X})$. PGMs are particularly of interest given their inherent explainability: each edge $(E_i \in \mathbf{E})$ $X \to Y$ in $\mathcal{G}$ is a graphical representation of the relationship between X and Y. This semantic interpretation allows researchers to gain high-level overviews of complex systems without sacrificing the direct connection with $P(\mathbf{X})$. A *Bayesian Network* (BN) [26] $\mathcal{B} = (\mathcal{G}, \Theta)$ with directed acyclic graph (DAG) $\mathcal{G}$ and parameters Θ is a PGM where $P(\mathbf{X})$ factorizes into local probability distributions according to $\mathcal{G}$ as $P(\mathbf{X}) = \prod_{X \in \mathbf{X}} P(X \mid Pa(X))$, where $Pa(X)$ is the parent set of X w.r.t. $\mathcal{G}$. An interesting property of $\mathcal{B}$ is that it encodes the *independence statements* in $P(\mathbf{X})$ into the DAG $\mathcal{G}$. Hence, it is possible to know whether X and Y are *probabilistically* independent in $P(\mathbf{X})$ by *graphically* querying $\mathcal{G}$, that is, one might ask if $X \perp\!\!\!\perp_P Y$ holds true by visiting $\mathcal{G}$, i.e. if X is independent of Y.

While representing probability independencies is useful to convey statistical associations, it is certainly more interesting to express causal-effect relationships to enable decision making, e.g., *rank_position* is a cause of *clickout*, therefore, intervening on the *rank_position* affects the *clickout*, but not the other way round. Under this representation, $\mathcal{G}$ is said to be a *causal graph* (CG) [27] and each edge $X \to Y$ in $\mathcal{G}$ is a *causal edge*, where X is said to be a *direct cause* of Y and Y is a *direct effect* of X. To be able to express more complex settings other than isolated edges, we need to formally represent the mechanism that generates the data that we observe, i.e., the *causal mechanism*. Again, we can leverage the CG $\mathcal{G}$ by defining a function f_X that assigns the value of X depending on the parents $Pa(X)$, so that $X := f_X(Pa(X), U_X)$, with U_X a random noise variable which accounts for non-deterministic interactions. When we superimpose the causal edge assumption on a BN we obtain a *Causal Network*, which encodes both probabilistic and causal interpretations.

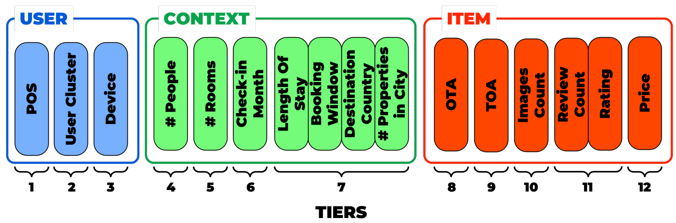

Figure 1: Edges tiers coloured w.r.t. their semantic interpretations. Factors are ordered according to the domain expert's prior knowledge, where a factor in lower tier can not be a cause of factor in a higher tier. For example, the destination country can not be a cause of the user's POS.

To construct a CN one must recover its CG, a process called *Causal Discovery* [28]. In this work, we relied on the *Hill-Climbing* (HC) [29] algorithm, which traverses the space of the possible CGs selecting the optimal graph $\mathcal{G}^*$ w.r.t. a goodness-of-fit function $\mathcal{S}$, known as the *scoring criterion*. At its core, HC iteratively modifies the current recovered graph to maximize $\mathcal{S}$ by adding, deleting or reversing individual edges. When no modification improves the score, the procedure halts and returns the current solution. HC is guaranteed to include edges that are coherent with the underlying independence statements, provided that $\mathcal{S}$ is a *consistent scoring criterion*. Another crucial aspect is the inclusion of *prior knowledge*. To this extent, domain experts can list specific edges that HC must exclude or include during the construction of the recovered graph, i.e., *forbidden* and *required edges*. Alternatively, it is also possible to define ordered edges sets, or *tiers*, that induce a partial order among the observed variables. This last encoding is particularly useful when (partial) temporal order of variables is known.

In this work, the hierarchy between user, context and items factors is defined by the RS and reported in Figure 1. Tiers are coloured w.r.t. their semantic interpretations, that is, users features are in blue, context features are in green and items features are in red. The result of the application of HC to the described data is reported in Figure 2, with the same colour scheme of Figure 1. We leveraged the independence statements to restrict the CG to a proper sub-graph by discarding irrelevant factors that do not affect, neither directly nor indirectly, the CTR.

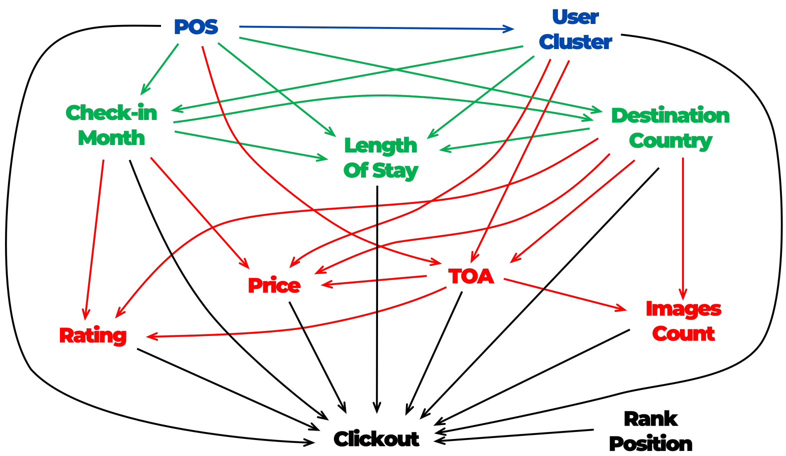

Figure 2: CG obtained by the causal discovery procedure with data and prior knowledge. Nodes are coloured by their semantic interpretation and edges are coloured according to the destination node.

4. Data Analysis Through Inference

In this section, we report a part of the data analysis performed using the CN presented in the previous section. Through this analysis, we want to find out how the factors affect user decision-making and evaluate the prediction of the model using expert knowledge. Therefore, we query the CN in the form $P(Y|rank\_position, X, W)$, where Y is the CTR, X is one item's feature and W is one user or context feature. The semantic interpretation of these queries is: "Which is the CTR for a certain *rank_position* if the user/context feature W had value w and the item's feature X had value x?". From this type of query, we can draw interesting conclusions like: "Given that the user/context W has value w, which is the value x of the item feature X that maximises the CTR Y?".

In our analysis, we submitted queries on all the combinations of one user/context feature and one item's feature in the CG of Figure 2, for a total of $5 \times 4 = 20$ combinations. We could only take into account one item's factor, one user/context's factor and the rank position since we can plot a limited number of dimensions. However, an RS is expected to perform similar reasoning to decide which items to recommend to a user in a given context, and it is not limited to three dimensions. In the following, we present two of the most interesting results to give an insight of our analysis. As previously shown in [16, 17, 18], the rank position has an important effect, regardless of other factors, since the CTR is higher for the first position and decreases for lower rank positions in both Figure 3 and Figure 4.

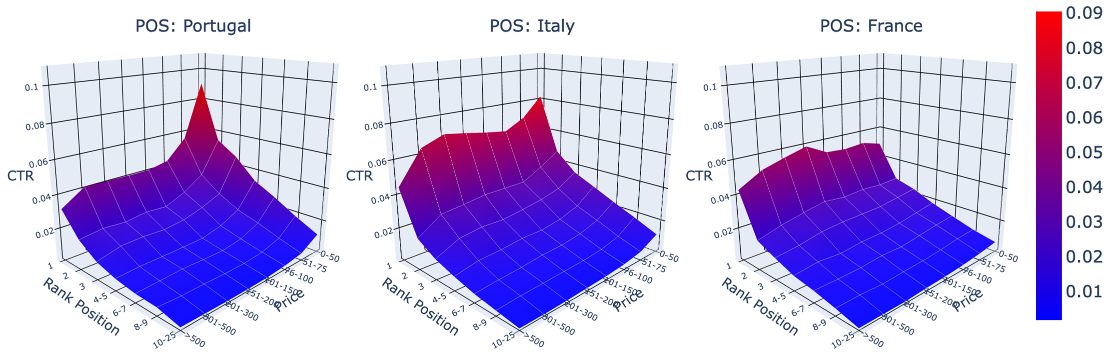

Figure 3: Results for the query $P(Y|rank\_position, POS, price)$ for POS of Portugal, Italy and France.

Figure 3 shows the results for the queries $P(Y|rank\_position, POS, price)$ for all the values of *price*, *rank_position* and three values of *POS* (Portugal, Italy and France). Different *POS*, which could be a proxy for the user's country, present different price sensitivities. Indeed, when $POS = Portugal$ (Figure 3, left), the CTR for every rank position is much higher for $price =$" $0 - 50$" than for higher prices. Instead, if we look at $POS = Italy$ (Figure 3, centre) this difference is much smaller, while for $POS = France$ (Figure 3, right), the CTR for $price =$" $0 - 50$" has the second smallest value after $price =$"> 500". From this analysis, we conclude that for $POS = Portugal$ we should recommend economic properties, while for $POS = France$, we must recommend more expensive properties that could lead to a higher profit without affecting the CTR (indeed, expensive properties were even more clicked in this case).

In Figure 4, we report the results of the query $P(Y|rank\_position, length\_of\_stay, TOA)$ where we can observe the CTR for different $length\_of\_stay$ and TOA. First, we note that long stays ("$9 - 14$ *nights*") have a lower CTR than shorter stays. However, it is usually not possible for the RS to change the user or the context, and thus, we cannot choose another $length\_of\_stay$ with a higher CTR. However, we can recommend different types of accommodations given different $length\_of\_stay$. Indeed, for "2 *nights*" (Figure 4, left), the highest CTR is associated with *1/2-star hotels (H-1/2*)*, *Guest Houses (GH)* and *Bed & Breakfast (B&B)*. Instead, for "$9 - 14$ *nights*" (Figure 4, right), the best *TOAs* (the ones with the highest CTR) are *4-star hotels (H-4*)*, *3-star hotels (H-3*)*, *Other TOAs (Otr)*, *Apartments (Apt)* and *Bed & Breakfast (B&B)*. In conclusion, for short stays, it is better to recommend "low quality" hotels or "temporary" solutions like B&B or Guest House, while, as the night of stay increases, the users prefer "higher quality hotels" or Apartments. For example, the *4-star hotels* are not the favourite option in case of 2 nights of

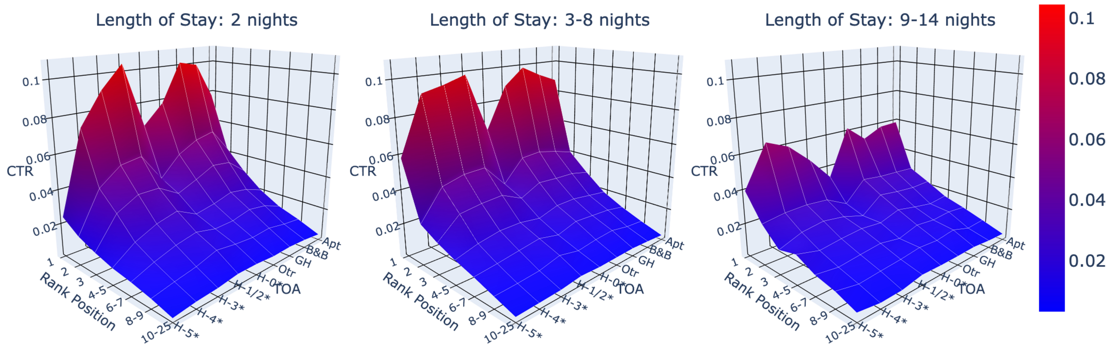

Figure 4: Results for query $P(Y|rank\_position, length\_of\_stay, TOA)$ for $length\_of\_stay$ of 2 nights, 3-8 nights and 9-14 nights.

stay, while are a good option for 3 to 8 nights of stay (Figure 4, centre) and are the best option for 9 to 14 nights of stay.

5. Conclusions

In this paper, we studied how different factors influence user click behaviour in online hotel searches. Specifically, we learned a CG by combining observational data provided by a meta-search booking platform with prior knowledge made available by domain experts. Then, based on the learned model, we performed an analysis to assess the influence of different factors on the user's decision to click or not on the recommended properties.

We assessed that CNs could be a valid estimator of user preferences. The analysis reported in this work matches our previous findings [16] where we also assessed the influence of ranking position and price on user decisions. As expected, the ranking position has a strong influence on CTR regardless of other factors and the price was confirmed to be a key factor influencing user's click behaviour as established from previous works, such as *Lockyer* [19] and *Stávková et al.* [20]. Moreover, the use of such a model allowed us to perform a more complex analysis, in which we could take into account several factors at once without losing explainability and semantic information. For example, it was also shown that the user market (POS) changed the influence of price as, for example, for POS equal to Portugal lower prices had the highest CTR, while for POS equal to France this was not true. Moreover, results suggest that a different number of nights changes user preferences in terms of the main types of accommodations. This work consequently highlights the many factors affecting users decision-making when performing online travel searches that are almost ignored in most offline studies.

Finally, the learned CG, together with the other necessary assumptions, opens the possibility of exploiting the methods developed in causality literature in the RSs domain. For example, it is possible to exploit a CG to analyse which factors really influenced user's decisions and model a RS accordingly. Furthermore, many approaches are developed in causality to account for confounding and selection bias. However, to apply these approaches a CG is needed to discover which factor we should adjust to obtain unbiased estimates. Finally, a CG enables also to estimate counterfactuals and opens to the possibility of generalising and transporting the inference made in one context to other contexts.

References

[1] F. Ricci, L. Rokach, B. Shapira, Introduction to recommender systems handbook, in: Recommender systems handbook, Springer, Boston, MA, 2011, pp. 1–35.

[2] D. Jannach, M. Zanker, A. Felfernig, G. Friedrich, Recommender systems: an introduction, Cambridge University Press, Cambridge, United Kingdom, 2010.

[3] J. Davidson, B. Liebald, J. Liu, P. Nandy, T. Van Vleet, U. Gargi, S. Gupta, Y. He, M. Lambert, B. Livingston, D. Sampath, The youtube video recommendation system, in: Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10, Association for Computing Machinery, New York, NY, USA, 2010, p. 293–296.

[4] C. A. Gomez-Uribe, N. Hunt, The netflix recommender system: Algorithms, business value, and innovation, ACM Trans. Manage. Inf. Syst. 6 (2016).

[5] Z. Yu, H. Xu, Z. Yang, B. Guo, Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints, IEEE Transactions on Human-Machine Systems 46 (2015) 151–158.

[6] W.-S. Yang, S.-Y. Hwang, itravel: A recommender system in mobile peer-to-peer environment, Journal of Systems and Software 86 (2013) 12–20.

[7] J. Zeng, F. Li, H. Liu, J. Wen, S. Hirokawa, 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), IEEE, 2016, pp. 55–60.

[8] D. Jannach, M. Zanker, M. Fuchs, Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations, J. Inf. Technol. Tour. 14 (2014) 119–149.

[9] M. Nilashi, O. bin Ibrahim, N. Ithnin, N. H. Sarmin, A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (em) and pca–anfis, Electronic Commerce Research and Applications 14 (2015) 542–562.

[10] M. P. O'Mahony, B. Smyth, Learning to recommend helpful hotel reviews, in: Proceedings of the Third ACM Conference on Recommender Systems, RecSys '09, Association for Computing Machinery, New York, NY, USA, 2009, p. 305–308.

[11] J. Srisuan, A. Hanskunatai, The ensemble of naïve bayes classifiers for hotel searching, in: 2014 International Computer Science and Engineering Conference (ICSEC), IEEE, Piscataway, New Jersey, United States, 2014, pp. 168–173.

[12] A. Levi, O. Mokryn, C. Diot, N. Taft, 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, RecSys '12, Association for Computing Machinery, New York, NY, USA, 2012, p. 115–122.

[13] M. Rossetti, F. Stella, M. Zanker, Analyzing user reviews in tourism with topic models, J. Inf. Technol. Tour. 16 (2016) 5–21.

[14] T.-Y. Liu, Learning to rank for information retrieval, Found. Trends Inf. Retr. 3 (2009) 225–331.

[15] M. Ludewig, D. Jannach, 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 '19, Association for Computing Machinery, New York, NY, USA, 2019, pp. 1–5.
[16] E. Cavenaghi, L. Camaione, P. Minasi, G. Sottocornola, F. Stella, M. Zanker, A re-rank algorithm for online hotel search, in: B. Ferrer-Rosell, D. Massimo, K. Berezina (Eds.), Information and Communication Technologies in Tourism 2023, Springer Nature Switzerland, Cham, 2023, pp. 53–64.
[17] E. A. Feigenbaum, H. A. Simon, A theory of the serial position effect, British Journal of Psychology 53 (1962) 307–320.
[18] T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, G. Gay, Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search, ACM Trans. Inf. Syst. 25 (2007) 7–es.
[19] T. Lockyer, The perceived importance of price as one hotel selection dimension, Tourism Management 26 (2005) 529–537.
[20] J. Stávková, L. Stejskal, Z. Toufarová, et al., Factors influencing consumer behaviour, ZEMEDELSKA EKONOMIKA-PRAHA- 54 (2008) 276.
[21] L. Coba, L. Rook, M. Zanker, Choosing between hotels: impact of bimodal rating summary statistics and maximizing behavioral tendency, Information Technology & Tourism 22 (2020) 167–186.
[22] J. Pearl, Causality, Cambridge university press, Cambridge, United Kingdom, 2009.
[23] M. A. Hernán, J. M. Robins, Causal Inference: What If, Boca Raton: Chapman \& Hall/CRC, United Kingdom, 2020.
[24] S. Xu, D. Xu, E. Korpeoglu, S. Kumar, S. Guo, K. Achan, Y. Zhang, Causal structure learning with recommendation system, arXiv preprint arXiv:2210.10256 1 (2022).
[25] D. Koller, N. Friedman, Probabilistic graphical models: principles and techniques, MIT press, Cambridge, Massachusetts, United States, 2009.
[26] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan kaufmann, Cambridge, Massachusetts, United States, 1988.
[27] M. Glymour, J. Pearl, N. P. Jewell, Causal inference in statistics: A primer, John Wiley & Sons, Hoboken, United States, 2016.
[28] A. Zanga, E. Ozkirimli, F. Stella, A survey on causal discovery: Theory and practice, International Journal of Approximate Reasoning 151 (2022) 101–129.
[29] M. Scutari, C. Vitolo, A. Tucker, Learning bayesian networks from big data with greedy search: computational complexity and efficient implementation, Statistics and Computing 29 (2019) 1095–1108.