A Product Feature-Based User-Centric Ranking Model for E-Commerce Search

Lamjed Ben Jabeur[1(✉)], Laure Soulier[2], Lynda Tamine[1], and Paul Mousset[1]

[1] IRIT, Université de Toulouse, CNRS, UPS, 118 Route Narbonne, Toulouse, France
{jabeur,tamine,mousset}@irit.fr

[2] Sorbonne Universités, UPMC Univ Paris 06, LIP6 UMR 7606, 75005 Paris, France
laure.soulier@lip6.fr

Abstract. During the online shopping process, users search for interesting products in order to quickly access those that fit with their needs among a long tail of similar or closely related products. Our contribution addresses *head* queries that are frequently submitted on e-commerce Web sites. Head queries usually target featured products with several variations, accessories, and complementary products. We present in this paper a product feature-based user-centric model for product search involving, in addition to product characteristics, the user engagement toward the product. This model has been evaluated through the product search track of the LL4IR lab at CLEF 2015 in order to highlight the effectiveness of our model as well as the impact of the user engagement factor.

Keywords: Information retrieval · Living labs · LL4IR · e-commerce · Product search · Ranking · User engagement · User preferences

1 Introduction

In the last few years, online retailers and marketplaces have shown a steady growth in terms of popularity as well as benefits. Amazon claims more than 240 million products available for sale on US store amazon.com[1]. The marketplace leader claims also more than 2 billion items sold worldwide by the end of 2014[2]. As the result of this huge quantity of available products, users are facing difficulty to make choice. The diversity of products in terms of functionalities and features makes their shopping experience more difficult.

To tackle this problem, online retailers enhance their Web sites with product search tools. In fact, product search is becoming more important [18], leading to propose adapted retrieval tools in order to help customers to find their products of interest [5]. One example of product search tools is proposed by Google Shopping for which customers have found the utility with around 100 billions of submitted queries by month[3].

---

[1] http://www.ecommercebytes.com/cab/abn/y14/m07/i15/s04.

[2] http://www.businesswire.com/news/home/20150105005186/en/Amazon-Sellers-Sold-Record-Setting-2-Billion-Items.

[3] http://www.godatafeed.com/resources/google-shopping-campaigns.

N. Fuhr et al. (Eds.): CLEF 2016, LNCS 9822, pp. 174–186, 2016.
DOI: 10.1007/978-3-319-44564-9_14

In the literature, product search has been addressed as an information retrieval (IR) task bridging e-commerce data and customer's information need formulated during the online shopping process. Previous works have proposed to integrate several features which might be split into two categories. On one hand, the proposed approaches focus mainly on the product fields, namely its category and its description [4,18]. On the other hand, users' preferences and search intent are emphasized leading to a user-centered search process [9].

In this paper, we propose a feature-based user-centric ranking model for product search that addresses the problem of head queries on e-commerce Web sites. Head queries represent the set of most frequent queries on featured products [1,17], such as dolls, miniatures, puzzles, cards. The latter may be characterized by several variations, accessories, and/or complementary products. Combining both approaches (product features [4,18] and user-centered [9]), our model ranks products with respect to their descriptive fields and category as well as their popularity highlighting the user engagement toward the product [14]. We evaluate our model while participating to the product search track of the Living Labs for IR (LL4IR) [17] of CLEF 2015 [11] and present also some analysis aiming at understanding the user engagement factor. More particularly, the contribution of our paper is twofold:

– A new product search model including both product characteristics and user engagement. This model is evaluated through the living lab paradigm.
– A statistical analysis highlighting how could be characterized the user engagement in terms of product search.

In the remainder of the paper, Sect. 2 synthesizes the state of the art surrounding product search. Section 3 introduces our product search model and describes its experimental evaluation. We present in Sect. 4 a statistical analysis on the effect of the user engagement on product search. Finally, Sect. 5 concludes the paper and presents perspectives.

2 Related Work

Similarly to the information access perspective, the literature review outlines several dimensions underlying the product search field. Some work focused on the understanding of the product search process according to the users' modeling perspective. First, Detlor et al. [8] compared the exploration and the search processes on e-commerce sites and outlined that the main difference relies on the type of users' intent with respect to the product specificity. More particularly, product search requires basic information (such as the price, the product description, or the information about the seller) as well as more complex specifications of a product. Second, other authors focused on interactions issued from film recommendation systems [2,3]. Although the tracked products (films) are different than the ones tracked by head queries of the LL4IR Labs, the authors highlighted that the diversity of recommended products is an important criterion in terms of users' satisfaction [2] and that it should be personalized to each user with respect to their past actions [2]. Moreover, the integration of the temporal diversity enables to avoid recommending redundant products retrieved over time [13] as well as to enable distinguishing short- and long-term preferences [19].

Other work, more close to our contribution, addressed the product search issue as an information retrieval challenge aiming at leveraging e-commerce data in order to answer customers' information needs.

The first line of work in this category includes retrieval models mainly based on product characteristics (e.g., the category and the description). Chen et al. [4] proposed to diversify product results and to return, among the large collection of similar products, only those significantly different from each others. Product categories and attribute values are used to diversify the list of products. Vandic et al. [18] addressed the issues underlying different hierarchical classifications in online stores and the multiple vocabulary terms used to describe the same product. Based on semantic ontologies, they proposed to match similar products and classify them into a universal product category taxonomy.

In the second line of work in the same category, the focus is oriented towards the user with an attempt to bridge the gap between the vocabulary used to describe the product and the customers' vocabulary used to formulate their search queries. For instance, the query *"cheap PC gamer"* might be difficult to solve by only comparing the query text with the product description since it requires reasoning over the search intent towards a particular product feature, namely the price. To tackle this challenge, Duan et al. [9] propose to represent both products and users through an entity-based representation in which each entity is formalized as a pair of key-value. The product retrieval is then performed through a probabilistic model which estimates the relevance at the level of attribute preferences. Other work leveraged users' search history in order to capture users' interests [16]. This type of model could be enhanced by product characteristics as done by Ghirmatsion and Balog [10] which proposed a model aiming at first identifying relevant products and then re-ranking products using relevance judgments of the search history. This approach has been enhanced by filtering techniques applied on product availability or reduction rate criteria.

In our contribution, we propose to combine both product and user point of view by *(1)* including product characteristics as previously done by [4,18] and *(2)* a metric highlighting the user engagement [14]. In contrast to [9,10,16] which focused on the interest of a particular user, our proposed user engagement metric leverages from the crowd.

3 Product Feature-Based User-Centric Ranking Model

In this section, we present our first contribution consisting in proposing and evaluating a product search model relying on a product feature-based user-centric approach. In the remaining section, we first present the model and then detail the experimental evaluation which has been carried out through the LL4IR lab.

3.1 The Model

Our model aims at leveraging product characteristics and user engagement towards the product. To do so, we estimate the relevance of product p with respect to query q as a combination of two indicators expressing the relevance probability $P(p|q)$ of product p based on its characteristic and a user engagement metric UE(p). The relevance $RSV(p, q)$ of product p given query q is computed as:

$$RSV(p, q) = P(p|q) * \mathrm{UE}(p) \tag{1}$$

The Product Feature-Based Probability. Products are commonly described in e-commerce Web site with multiple fields[4]. These fields enable to identify the product (i.e., sku, gtin13, ISBN), describe its purpose (i.e., name, brand, description), list elementary and technical features (i.e., model, speed, weight, color) as well as organize the product collection into a structured hierarchy (i.e., category). With this in mind and inspired by work of Craswell et al. [6] and Dakka et al. [7], we propose to depict product p in two sets of elements consisting in (a) its set of textual descriptive fields d_p that describe the product, and (b) its category that organizes products by categories.

Accordingly, the relevance $P(p|q)$ of product p with respect to query q could be rewritten as $P(c_p, d_p|q)$ (Eq. 2). According to Bayes probability rules (Eq. 3) and assuming that the product category and description are independent (Eq. 4), product relevance is estimated by the following model:

$$P(p|q) = P(c_p, d_p|q) \tag{2}$$
$$= P(c_p|q) \cdot P(d_p|c_p, q) \tag{3}$$
$$\propto P(c_p|q) \cdot P(d_p|q) \tag{4}$$

where $P(c_p|q)$ and $P(d_p|q)$ express respectively the relevance of category c_p of product p and the topical relevance of product description d_p with respect to query q. We detail these probabilities below.

- *Topical relevance of product description* d_p. The topical relevance focuses on the product descriptive field set d_p. Except for the category field, all remaining fields are part of the product description d_p. We consider *(1)* the title which is usually size limited and includes concise information about the product and *(2)* the description field including broader information.

We propose to use the $BM25F$ scoring schema [6,20] to estimate likelihood $p(d_p|q)$ of descriptive fields d_p given query q. The $BM25F$ computes the similarity with query q while attributing different weights to each field.

We first calculate normalized term frequency $\overline{tf}_{t,f,p}$ for each field:

$$\overline{tf}_{t,f,p} = \frac{tf_{t,f,p}}{1 + b_f(\frac{l_{f,p}}{l_f} - 1)} \tag{5}$$

[4] http://schema.org.

where $tf_{t,f,p}$ represents the frequency of term t in field f belonging to description d_p of product p. $l_{f,p}$ is the length of field f in product description d_p and l_f is the average length of field f, b_f is a field-dependant parameter similar to b parameter in $BM25$. The term frequencies estimated over all the field set are combined linearly using weight w_f of field f as follows:

$$\overline{tf}_{t,p} = \sum_{f \in d_p} w_f * \overline{tf}_{t,f,p} \tag{6}$$

The term frequency $\overline{tf}_{t,p}$ is integrated in the usual $BM25$ saturating function modeling the non-linear relevance distribution of term frequencies. The probability $p(d_p|q)$ is approximated by the $BM25F$ function [6,20]:

$$p(d_p|q) \approx BM25F(q|d_p) = \sum_{t \in q \cap d_p} \frac{\overline{tf}_{t,p}}{k_1 + \overline{tf}_{t,p}} idf(t) \tag{7}$$

where k_1 and $idf(t)$ express respectively the BM25 parameter and the inverse document frequency of term t.

- *The relevance of category.* The relevance of category c_p with respect to query q aims at identifying to what extent the category is relevant to the product collection. The underlying idea is to decide which eminent category likely matches the query since different categories may respond to the query.

Let S be the set of non-negative topical scores obtained by product description d_p of all products $p \in D(c_p)$, where $D(c_p)$ corresponds to the set of product characterized by category c_p. More formally, S is defined as follows:

$$S = \{p(d_p|q)|p \in D(c_p) \wedge p(d_p|q) \geq 0\} \tag{8}$$

where $p(d_p|q)$ is approximated by $BM25F(q,d)$ as done in Eq. 7. We propose to estimate the relevance likelihood $p(c_p|q)$ of product category c_p towards query q with similarity $sim(q,c_p)$ of product category c_p given query q. This similarity is estimated as the product of the log scale cardinality of set S and an aggregation function $A(S)$ of topical scores over respective products:

$$p(c_p|q) \approx sim(q,c_p) = log(1 + |S|) * \mathcal{A}(S) \tag{9}$$

where $\mathcal{A}(S)$ can be computed as the maximum, the mean and the median scores over the topical distribution of all products $D(c_p)$. We propose to use the 95^{th} percentile as aggregate function $\mathcal{A}(S)$. In contrast of mean and maximum, the 95^{th} percentile is resistant to outliers. Similarly to the median, 95^{th} percentile allows measuring the global tendency of topical scores.

As the category includes more relevant products with respect to the query, the category might be relevant to the query. This is reflected by the first part of Eq. 9, noted $log(1 + |S|)|$. The log scale value enables to lower high cardinality and thus smooths the importance of overpopulated categories.

The User Engagement Metric. The integration of the user engagement component is driven by the main aim of e-commerce application which consists in increasing the user conversion rate. Although the user engagement should be derived in accordance with the application goal, such metric emphasizes the positive aspect of the interaction [14]. In the setting of a Web application, the user's engagement is often associated with his/her interactions including visits, clicks, comments, recommendations, etc. In accordance with the search scenario of this model, we propose to consider users' interactions willing to be noticed after a product search. For instance, post-task evidence sources of interaction could be result clicks, product ratings, favorites, or users' actions. The latter aim at bookmarking wishlist, adding to basket, and/or pushing the product. Unfortunately, these data are not available for this edition of LL4IR track. Thus, we estimate the user engagement by the number of social interactions, namely "Like" and "Share" actions, generated on the Facebook[5] social media platform.

In order to get the social engagement toward a product, we first identify significant Web resources that represent a product, typically Web pages with technical description. In this aim, we used the product name as a query for exact search on a Web search engine. We assume that the set of top k resources significantly represent the product and their underlying users' interactions may be associated with the product. With this in mind, let $R_p = \{r_1, r_1 \cdots r_k\}$ be the set of resources that mention product p. $likes(r_i)$, respectively $shares(r_i)$, expresses the number of Facebook likes, respectively Facebook shares, obtained by a particular resource r_i. Please, note that likes and shares are obtained by sending the URL of resource r_i to the Facebook API.

The user engagement of resource r_i identified through is computed is follows:

$$e(r_i) = \frac{\log(1 + \min(s, likes(r_i) + shares(r_i)))}{\log(1 + s)} \tag{10}$$

with s defines an upper bound on social interactions of resource r_i.

In the end, the user engagement of product p corresponds to the maximal user engagement obtained by the associated resources:

$$\mathrm{UE}(p) = \mathrm{argmax}_{r_i \in R_p}\, e(r_i) \tag{11}$$

3.2 Experimental Evaluation

In order to evaluate the effectiveness of the proposed model, we relied on the "Living Labs for Information Retrieval" (LL4IR) campaign [17] aiming at evaluating IR models in real utilization's cases: users submit their queries on a website and interact in real time with retrieved results of participants. The evaluation campaign proposes several evaluation periods (also called "rounds") in which the main search task consists in a product search task on the online commerce site of REGIO JTK[6], Hungarian leader in the sale of toy for children. In this section,

[5] http://facebook.com.
[6] http://www.regiojatek.hu/.

we first describe the experimental context implemented during the LL4IR campaign, by introducing the protocol design and the obtained results.

Protocol Design. *Experimental data:* The LL4IR campaign provides a set of experimental data:

- 100 oriented product queries extracted from most frequent queries submitted on the system in the past. To allow comparability between rounds, queries are the same over all rounds. We note that half of the queries are used for training.
- A product collection including both available products and those labeled as unavailable which would be available later. The average number of products associated with each query is around 60 products. Each product is represented by a set of structural and semantic meta-data, like the characters associated with the product (e.g. Spiderman, Hello Kitty), it brand (e.g. Beados, LEGO), or the recommended age/gender.
- The user feedback updated every 5 min throughout a specific round. Each user feedback is represented by a binary value, depending on whether the product presented was clicked by the user.

Evaluation protocol. The aim of the LL4IR campaign is to leverage users' clicks in order to compare the effectiveness of the systems proposed by the participants with respect to the one of the production system. To do so, product ranking of each participating system is interleaved with the product ranking of the production system. The latter corresponds to the default product ranking system provided by Web site owners. For each submitted query belonging to the preselected head query set, the user gets a set of results for which the half comes from website production system and the other half from a random participating system. The same process has been carried out over a baseline model proposed by the organizers of the campaign [17] and other participants [10,16].

Metrics. Five metrics, estimated over all submitted head queries, are proposed by Living Labs organizers in order to evaluate a participating system:

- The number of wins, noted $\#Wins$, which expresses the number of times the test system received respectively more clicks than the product system.
- The number of losses, noted $\#Losses$, which expresses the number of times the test system received respectively fewer clicks than the product systems.
- The number of ties, noted $\#Ties$, which expresses the number of times the test system received respectively as many clicks as the product systems.
- The number of Impressions, noted $\#Impressions$, which expresses the number of times the test system is mixed with production one with $\#Impressions = \#Wins + \#Losses + \#Ties$
- The outcome, noted *Outcome*, is defined as the ratio of wins over the sum of wins and losses (Eq. 12). A ratio higher than 0.5 highlights the system ability

to provide more relevant products than irrelevant ones, assuming that clicks are indicators of product relevance [12].

$$Outcome = \frac{\#Wins}{\#Wins + \#Losses} \tag{12}$$

Results. In order to evaluate both components of our model, we tested our model differently over rounds (Round 2 and 3 - since we did not participate to round 1): for round 2, we only rank products according to the characteristic-based indicator using Eq. 4 while for round 3 we introduced the user engagement-based indicator as explained in Eq. 1. We outline that, since the LL4IR campaign allows participants to submit only one run for each testing period, we fixed the descriptor and parameter weights according to previous work, respectively [20] for descriptor and [6,20] for the BM25F parameters.

Table 1 presents results obtained by the baseline, the best concurrent participant and our model for these two rounds. We could see that for round 2 we obtained the lowest outcome measure with respect to the baseline and the participant. Results obtained for round 3 highlight that the user engagement allows enhancing the effectiveness of our model. Please note that this statement is limited since the evaluation metric might be impacted by the set of users involved in the evaluation process which is variable over the different rounds. We outline that the ranking model of the system, the queries and the interleaved method are stable over the different rounds.

Table 1. Effectiveness comparison of our model during round 2 and 3 of the LL4IR

	Round 2		Round 3	
	Outcome	% Chg	Outcome	% Chg
Baseline	0.5284	-24.48 %	0.4430	+10.38 %
Best participant	0.4795	-16.78 %	0.4507	+8.49 %
Our model	0.3990		0.4890	

However, the comparison with the baseline as well as the best participant highlights that our model obtains the highest outcome value (0.489) with improvements higher than +8.49 %. This reinforces our intuition that the user engagement should be integrated into IR models [14]. Moreover, our model obtains an outcome value for round 3 closed to 0.5, suggesting that its effectiveness is relatively similar to the one of the product search model of the e-commerce website. The outcome values obtained by the participants are generally even more lower than 0.5. Taken in a whole, this results show the difficulty of formalizing retrieval models for product search, which is a quite young research domain.

In order to compare the effectiveness at the query level, we plot in Fig. 1 results of our model at the query level for round 2 and 3, highlighting the impact

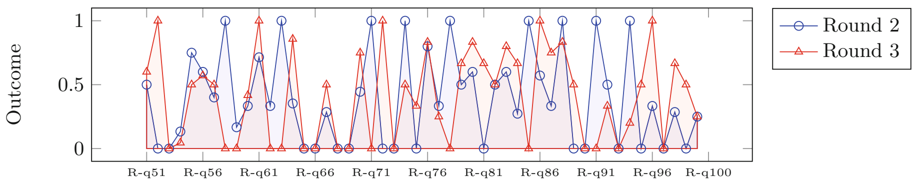

Fig. 1. Effectiveness comparison of our model between round 2 and 3 - Impact of the user engagement metric.

of the user engagement. A descriptive analysis shows that the user engagement indicator enables to improve the effectiveness of 23 queries over the 50 ones, with an average difference equals to 0.391 for those 23 queries (against 0.083 for all queries). Accordingly, we hypothesize that it exists two types of head queries depending on whether they benefit from user engagement factor or not in terms of retrieval effectiveness. In the remaining paper, we call "socially-motivated queries" those leveraging user engagement. A quick overview of query text emphasizes that "socially-motivated queries" seem to be those expressing non-targeted and specific products (e.g., "puzzle", "doll house" or "ball"). In contrast, the "non-socially-motivated queries" (with null or negative improvements) more particularly refer to focused products, generally addressed by a brand (e.g., "Playmobil", "Cars", "Scrabble" or "Angry birds"). This results lead us to analyze more in-depth the user engagement factor.

4 Understanding the Effect of the User Engagement Factor in Product Search

In this section, we address the second contribution of our paper aiming at understanding the user engagement factor. In this aim, we propose to deepen our analysis with a twofold objective: *(1)* identifying "socially-motivated queries" characteristics and *(2)* highlighting the characteristics of product rankings associated to "socially-motivated queries"/"non-socially-motivated queries".

First, we consider two classes of queries (namely "socially-motivated queries" and "non-socially-motivated queries"). We performed a logistic regression aiming at explaining the social responsiveness of queries according to the query characteristics. The latter are those provided by the LL4IR platform (namely, the number of users' clicks, the number of products, the absence/presence of concepts - noted *has.concept*, the absence/presence of brand, the absence/presence of famous character - noted *has.character*) as well as estimated features, namely, the query length. In addition to these characteristics extracted through the textual analysis of the query, we propose to extract new ones based on search result clusters. In particular, we propose to use an unsupervised clustering algorithm called "Lingo" [15]. The latter applies phrases analysis and latent semantic techniques with the aim of clustering search results into meaningful groups. For each query, we then obtained the following features: the number of associated clusters

Table 2. Descriptive model of user preferences for product search on e-commerce sites

Characteristics	Regression estimate	p-value
$NbrClusters$	0.015	0.0139 *
$AvgClusterSize$	-0.574	0.0225 *
$MinClusterSize$	0.681	0.0106 *
$has.concept$	0.451	0.0222 *
$is.character$	-0.379	0.0500 *

Table 3. Descriptive model of user preferences for product search on e-commerce sites

Query class	Characteristics	Reg. estimate	p-value
"Socially-motivated queries"	Price	0.002	$\leq$ 2e-16 ***
	Gender (Male)	0.0044770	8.65e-02
	Gender (Female)	0.0165339	1.29e-04 ***
"Non-socially-motivated queries"	Number of pictures	0.008	9.81e-08 ***
	Discount rate	-0.009	6.74e-03 **
	Bonus	0.0329	2.59e-03 **

(*NbrClusters*), the size of the largest cluster (*MaxClusterSize*), the average size of clusters (*AvgClusterSize*), and the minimal size of clusters (*MinClusterSize*).

At each iteration of the backward method, we removed the product characteristic with the highest p-value until all characteristics involved within the model impact significantly on the class ($p-value \leq 0.5$). A positive and significant regression estimate of a particular feature expresses the fact that the higher the value of the feature, the more the query is "socially-motivated". The final statistical model with significant features is presented in Table 2. Results suggest that "socially-motivated queries" are generally queries not referring to famous characters but rather expressing a concept typically related to main themes of products (e.g. "guitar", "kitchen", etc.). Also, those queries generally lead to diversified products. This is shown through the positive correlation with small clusters since the obtained clusters contain few products.

Second, we propose to analyze users' product preferences for both types of queries. We believe that such analysis would help the community to build more effective models for this particular application domain once they have identified "(non-)socially-motivated queries". In this aim, for each query, we consider as evidence source the whole set of products provided by the LL4IR campaign. Instead of using history of click rates that are highly correlated to time ("product trend") and the product availability, we infer the users' preferences from a metric provided by the LL4IR organizers expressing the probability of a product to be clicked by a user. This probability is estimated for a round by the ratio of clicks received by a product and the number of times the product was presented to the user. Accordingly, we build the statistical model for each query class aiming at identifying users' preferences with respect to product characteristics. We used a generalized model and, as done earlier, at each iteration, we removed the product characteristic with the highest p-value until all characteristics impact significantly on the click-based measure ($p - value \leq 0.5$). The higher the value of the feature, the higher the product probability being clicked is. The results are presented in Table 3.

One can see that different features characterize the two query classes, but the price of the product seems to be an important decision-making factor. Indeed, users submitting "socially-motivated queries" are generally interested in products with a higher price than those submitting "non-socially-motivated queries". In addition, the latter users appreciate products with a discount rate (*Bonus*) although low (negative regression estimate of *Discount rate*). Coupled with the query characteristics analysis, this suggests that users expressing "socially-motivated queries"(which mainly address non-targeted information need in terms of brands and characters) considers the price as a product quality indicator. The latter should allow users distinguishing similar products (e.g., among the different types of baby dolls). In contrast, users addressing" non-socially-motivated queries" are looking for particular products with specific characteristics of brands and characters and accordingly appreciate less expensive products.

The gender seems to be an important factor for "socially-motivated queries", orienting the product search model towards products for female. However, it is difficult to infer strong statements from this feature since we do not know the population of users submitting those queries.

Last, the descriptive model reveals that "non-socially-motivated queries" require a picture while this factor is not discriminant for "socially-motivated queries". This suggests that users expressing non-targeted information needs remain general in their decision-making and express small requirements, excepting the price. However, users formulating product need towards specific brands, concepts and characters stay focused on the product design, and the picture is a way to capture the specificity and the credibility of the product. In this case, the presence of pictures is a triggering purchase factor, which is already well-known as a marketing strategy highlighted by some studies. The latter reveals that 67 % of consumers considers product pictures as extremely important[7].

5 Discussion and Conclusion

We presented a product feature-based user-centric model for product search involving in addition to product characteristics the user engagement toward the product. The experimental evaluation has been carried out through the LL4IR framework and suggests that the user engagement is an interesting factor in

[7] http://blog.lemonstand.com/7-ways-optimize-product-page-conversions/.

product search. To better understand this factor, we performed statistical analysis highlighting characteristics of queries. With respect to a query classification based on the user engagement responsiveness, we also identify users' preferences in terms of products. These results are not without limitations since analysis are dependent of the experimental framework biases. However, we believe that the naturalness of the experimental evaluation allows considering these results as reasonable. Moreover, our estimation of the user engagement relies on the product popularity [14], but should be refined according to users' interactions.

From this analysis, we pointed out interested users' behaviors and preferences in terms of product search that could be useful for the design of retrieval models in this application domain. One particular statement revealed that some queries are sensible to the user engagement, impacting the features of the product ranking algorithms. For instance, the price is a pivotal feature in product search which should be used differently according to the users' need (users expressing "socially-motivated queries" seems to be willing to buy more expensive products than those expressing "non-socially-motivated queries"). In the future, we plan to enhance our product search model to take into consideration the findings of this paper by proposing a query-adapted product search models which *(1)* detects the query type by taking into account their categories or whether it implies concepts or famous characters, and *(2)* rank products using features that particularly attract users.

References

1. Balog, K., Kelly, L., Schuth, A.: Head first: living labs for ad-hoc search evaluation. In: CIKM 2014, pp. 1815–1818. ACM (2014)
2. Castagnos, S., Jones, N., Pu, P.: Eye-tracking product recommenders' usage. In: RecSys 2010. ACM (2010)
3. Chen, L.: Social influence of product popularity on consumer decisions:usability study of flickr camera finder. In: IUI 2010, pp. 297–300 (2010)
4. Chen, X., Wang, H., Sun, X., Pan, J., Yu, Y.: Diversifying product search results. In: SIGIR 2011, pp. 1093–1094. ACM (2011)
5. Corbitt, B.J., Thanasankit, T., Yi, H.: Trust and e-commerce: a study of consumer perceptions. Electron. Commer. Res. Appl. **2**(3), 203–215 (2003)
6. Craswell, N., Zaragoza, H., Robertson, S.: Microsoft cambridge at TREC 14: enterprise track. In: TREC 2005 (2005)
7. Dakka, W., Gravano, L., Ipeirotis, P.G.: Answering general time-sensitive queries. IEEE ToKDE **24**, 220–235 (2012)
8. Detlor, B., Sproule, S., Gupta, C.: Pre-purchase online information seeking: search versus browse. J. Electron. Commer. Res. **4**(2), 72–84 (2003)
9. Duan, H., Zhai, C.X., Cheng, J., Gattani, A.: Supporting keyword search in product database: a probabilistic approach. Proc. VLDB Endow. **6**(14), 1786–1797 (2013)
10. Ghirmatsion, A.B., Balog, K.: Probabilistic field mapping for product search. In: CLEF 2015 Workshop (2015)
11. Mothe, J., Savoy, J., Kamps, J., Pinel-Sauvagnat, K., Jones, G., San Juan, E., Capellato, L., Ferro, N. (eds.): CLEF 2015. LNCS, vol. 9283. Springer, Heidelberg (2015)

12. Joachims, T.: Optimizing search engines using clickthrough data. In: KDD 2002, pp. 133–142 (2002)
13. Lathia, N., Hailes, S., Capra, L., Amatriain, X.: Temporal diversity in recommender systems. In: SIGIR 2010, pp. 210–217. ACM (2010)
14. Lehmann, J., Lalmas, M., Yom-Tov, E., Dupret, G.: Models of user engagement. In: Masthoff, J., Mobasher, B., Desmarais, M.C., Nkambou, R. (eds.) UMAP 2012. LNCS, vol. 7379, pp. 164–175. Springer, Heidelberg (2012)
15. Osiński, S., Weiss, D.: A concept-driven algorithm for clustering search results. IEEE Intell. Syst. **20**(3), 48–54 (2005)
16. Schaer, P., Tavakolpoursaleh, N.: Historical clicks for product search: Gesis at clef ll4ir 2015. In: CLEF 2015 Workshop (2015)
17. Schuth, A., Balog, K., Kelly, L.: Overview of the living labs for information retrieval evaluation (LL4IR) CLEF lab 2015. In: Mothe, J., Savoy, J., Kamps, J., Pinel-Sauvagnat, K., Jones, G., San Juan, E., Capellato, L., Ferro, N. (eds.) CLEF 2015. LNCS, vol. 9283, pp. 484–496. Springer, Heidelberg (2015). doi:10.1007/978-3-319-24027-5_47
18. Vandic, D., van Dam, J.-W., Frasincar, F.: Faceted product search powered by the semantic web. Dec. Supp. Systems **53**(3), 425–437 (2012)
19. Xiang, L., Yuan, Q., Zhao, S., Chen, L., Zhang, X., Yang, Q., Sun, J.: Temporal recommendation on graphs via long- and short-term preferencefusion. In: KDD 2010, pp. 723–732. ACM(2010)
20. Zaragoza, H., Craswell, N., Taylor, M.J., Saria, S., Robertson, S.E.: Microsoft cambridge at TREC 13: web and hard tracks. In: TREC 2004 (2004)