UvA-ExPoSe at TREC 2016: Contextual Suggestion Track

Dehghani, M.; Azarbonyad, H.; Kamps, J.; Marx, M.

Publication date
2016

Document Version
Final published version

UvA-ExPoSe at TREC 2016: Contextual Suggestion Track

Mostafa Dehghani[1] Hosein Azarbonyad[2] Jaap Kamps[1] Maarten Marx[2]
[1]Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands
[2]Informatics Institute, University of Amsterdam, The Netherlands
{dehghani,h.azarbonyad,kamps,maartenmarx}@uva.nl

ABSTRACT

In this paper, we present the participation of the University of Amsterdam's ExPoSe team in the TREC 2016 Contextual Suggestion Track. The main goal of contextual suggestion track is to evaluate methods for providing suggestions for activities or points of interest to users in a specific location, at a specific time, taking their personal preferences into consideration. One of the key steps of contextual suggestion methods is estimating a proper model for representing different objects in the data like users and attractions. Here, we describe our approach which is employing Significant Words Language Models (SWLM) [1] as an effective method for estimating models representing significant features of sets of attractions as user profiles and sets of users as group profile.

We observe that using SWLM, we are able to better estimate a model representing the set of preferences positively rated by users as their profile, compared to the case we use standard language model as the profiling approach. We also find that using negatively rated attractions as negative samples along with positively rated attractions as positive samples, we may loose the performance when we use standard language model as the profiling approach. While, using SWLM, taking negatively rated attractions into consideration may help to improve the quality of suggestions. In addition, we investigate groups of users sharing a property (e.g. of a similar age) and study the effect of taking group-based profiles on the performance of suggestions provided for individual users. We noticed that group-based suggestion helps more when users have a tendency to rate attraction in a neutral way, compared to the case users are more subjective in their rating behavior.

Keywords

Contextual Suggestion, TREC, Significant Words Language Model

1. INTRODUCTION

This paper presents the participation of University of Amsterdam, ExPoSe team, in the TREC 2016 Contextual Suggestion Track. Contextual suggestion is the task of searching for complex information needs that are highly dependent on both context and user interests. More precisely, given the information of users including their age, gender, and set of rated places or activities as the user preferences (ratings are in the range of -1 to 4), the task is to generate a list of ranked suggestions from a set of candidate attractions, by giving the user information as well as some information about the context, including location of trip, trip season, trip type, trip duration, and the type of group the person is travelling with.

One of the key steps of the contextual suggestion methods is estimating a proper model for representing different objects in the data, like users and attractions. Having a proper profile of different objects representing significant features of them, we are able to find connections between data objects and provide effective suggestions based on them.

Our approach for TREC 2016 Contextual Suggestion Track is based on estimating effective profiles for different objects in the data. Generally speaking, we use Significant Words Language Models (SWLM) [1] as an effective method for estimating models representing significant features of a set of documents. We employed SWLM in different cases for modeling different objects in the contextual suggestion task, including: building a positive profile for each user (as a set of positively rated documents), building a negative profile for each user (as a set of negatively rated documents), building a profile for a group of users (as a set of users sharing a specific property) [2, 3].

In this paper, we address two main research questions:

RQ1 *How can SWLM help to estimate better user profiles for the contextual suggestion?*

RQ2 *In what conditions does information of group profiles estimated using SWLM improve the performance of the contextual suggestion?*

Briefly, we have made use of the estimated models by SWLM in different experiments with different settings to investigate how SWLM helps to improve the general performance of contextual suggestion. We observe that using SWLM we are able to better estimate a model representing the set of preferences positively rated by users as their profile, compared to the case we use standard language model as the profiling approach.

Furthermore, we find that using negatively rated attractions as negative samples along with positively rated attractions as positive samples, we may loose the performance when we use standard language model as the profiling approach. The reason is that standard language model of negatively rated documents contain positive information and using this model as the negative model, we penalize positive candidates. However employing SWLM we can estimate an effective model represent the essential negative terms which taking this model into consideration as the negative model, we are able to get rid of negative candidates and generally improve the of suggestions.

Moreover, we investigate the effect of employing information from groups that users belong to on the performance of suggestions provided for individual users. We observed that providing group-based suggestions helps to improve the performance of contextual suggestion for some users and it is not effective for others. Looking into the differences of these users, we noticed that for users that rate attractions mostly with rates close to the neutral rating, we improve the performance by taking group information into account, while for users that rate attraction in more subjective way, (i.e. giving high rates or low rates), group information do not help too much. The reason would be the fact that in the case of neutral rating behavior, we do not have strong signals and group-based information are counted as complementary signals helping the method to provide better suggestions.

2. ESTIMATING EFFECTIVE PROFILES

In this section, we explain how to estimate significant words language models and how to use them in contextual suggestion task.

2.1 Estimating SWLM

Having a set of documents, $\mathcal{S}$, in order to estimate SWLM representing the whole set, we assume that there are three models from which each document in the set is generated as a mixture sampling from these models: *significant words* model, *general* model, and *specific* model. The significant words model represents the latent model that is the distribution of terms reflecting the essential features of the object. The general model represents commonly observed terms and the specific model represents the partially observed terms in the set, which we assume as two different patterns of distribution of non-significant terms.

Each model is represented using a terms distribution, or a unigram language model, θ_{sw}, θ_g, and θ_s. Based on the generative model, each term in a document in the set, is generated by sampling from a mixture of these three models independently. Thus, the probability of appearance of the term t in the document d is as follows:

$$p(t|d) = \lambda_{d,sw} p(t|\theta_{sw}) + \lambda_{d,g} p(t|\theta_g) + \lambda_{d,s} p(t|\theta_s), \quad (1)$$

where $\lambda_{d,x}$ stands for $p(\theta_x|d)$ which is the probability of choosing the model θ_x given the document d.

We estimate θ_g and θ_s and make them fixed in the estimation process. We consider the collection model, θ_C as an estimation for θ_g:

$$p(t|\theta_g) = p(t|\theta_C) = \frac{c(t,C)}{\sum_{t' \in V} c(t',C)}, \quad (2)$$

where $c(t,C)$ is the frequency of term t in the collection. This way, terms that are well explained in the collection model get high probability and are considered as general terms.

Furthermore, we define specificity as being supported by part of documents in the set but not all of them. We estimate θ_s to represent the probability of a term being partially observed as follows, and rescale all the probabilities, to recover the probability values and establish a well-formed distribution:

$$p(t|\theta_s) \xleftarrow{Normalized} \sum_{d_i \in \mathcal{S}} \Bigg(p(t|\theta_{d_i}) \prod_{\substack{d_j \in \mathcal{S} \\ j \neq i}} (1 - p(t|\theta_{d_j})) \Bigg), \quad (3)$$

where $P(t|\theta_{d_i}) = {c(t,d_i)}/{\sum_{t' \in d_i} c(t',d_i)}$. Intuitively, Equation 3 calculates the probability of term t to be a specific term. To this end, it considers the probability of a term to be important in one of the document models but not others, marginalizing over all documents in the set. This way, terms that are well explained in only one document in the set but not others get higher probabilities and are considered as insignificant specific terms.

Having the above assumptions, the goal is to fit the log-likelihood model of generating all terms in the documents in the set to discover the term distribution of the significant words model, θ_{sw}. Let $\mathcal{S} = \{d_1, \ldots, d_{\mathcal{S}}\}$ be the set of documents. The log-likelihood function for the entire set of documents is:

$$\log p(\mathcal{S}|\Upsilon) = \sum_{d \in \mathcal{S}} \sum_{t \in V} c(t,d) \log \Big(\sum_{x \in \{sw,g,s\}} \lambda_{d,x} p(t|\theta_x) \Big), \quad (4)$$

where $c(t,d)$ is the frequency of the term t in the document d, and Υ determines the set of all parameters that should be estimated, $\Upsilon = \{\lambda_{d,sw}, \lambda_{d,g}, \lambda_{d,s}\}_{d \in \mathcal{S}} \cup \{\theta_{sw}\}$.

To fit our model, we estimate the parameters using the maximum likelihood (ML) estimator. Therefore, assuming that documents are represented by a multinomial distribution over the terms, we solve the following problem:

$$\Upsilon^* \coloneqq \operatorname*{argmax}_{\Upsilon} p(\mathcal{S}|\Upsilon) \quad (5)$$

Assuming that $X_{d,t} \in \{sw, g, s\}$ is a hidden variable indicating which model has been used to generate the term t in the document d, we can compute the parameters using the Expectation-Maximization (EM) algorithm. The stages of the EM algorithm are as follows:

E-Step

$$p(X_{d,t} = x) = \frac{p(\theta_x|d) p(t|\theta_x)}{\sum_{x' \in \{sw,g,s\}} p(\theta_{x'}|d) p(t|\theta_{x'})} \quad (6)$$

M-Step

$$p(t|\theta_{sw}) = \frac{\sum_{d \in \mathcal{S}} c(t,d) p(X_{d,t} = sw)}{\sum_{t' \in V} \sum_{d \in \mathcal{S}} c(t',d) p(X_{d,t'} = sw)} \quad (7)$$

$$\lambda_{d,x} = p(\theta_x|d) = \frac{\sum_{t \in V} c(t,d) p(X_{d,t} = x)}{\sum_{x' \in \{sw,g,s\}} \sum_{t \in V} c(t,d) p(X_{d,t} = x')} \quad (8)$$

As explained above, besides removing common terms by advocating terms that are relatively rare in the collection, the main contribution of SWLM is that it eliminates specific terms by favoring terms occurring in all the feedback documents, not only some of them, which leads to an effective language model as the profile of the object.

2.2 SWLM for Contextual Suggestion

As it is explained, SWLM is an effective approach to estimated a model representing a set of documents. In our setting, each profile, either for an individual user or for a group of users, is a set of textual documents and we use language models as the representation of profiles.

In the case of estimating a profile for an individual user, the set of documents contains textual information of attractions rated by the user is the input of SWLM. This is both in the case of estimating a positive user profile using the set of

Table 1: Performance of Contextual Suggestion Using Different User Profiling Approaches

Method	ndcg@5	recip_rank	P@5
$UserProfile^{+}_{SLM}$	0.1832	0.5015	0.3108
$UserProfile^{+/-}_{SLM}$	0.1301	0.4401	0.2312
$UserProfile^{+}_{SWLM}$	0.2672	0.5453	0.3548
$UserProfile^{+/-}_{SWLM}$	**0.2711**	**0.5453**	**0.3911**

positively rated attractions, and estimating a negative user profile using the set of negatively rated attractions.

With regard to the data of the TREC contextual suggestion task, we estimate user positive profile using SWLM on the set of documents that are rated higher than 2 (i.e 3 and 4) and estimate user negative profile using SWLM on the set of documents that are rated lower than 2 (i.e. 1 and 0, with ignoring documents with rate = -1). In order to take given rates by users into account in the process of SWLM, we add a pseudo-count for terms in documents with rate 4 in positive profiling and terms in documents with rate 0 in negative profiling. In other words, we update the value of $c(t, d)$ in the EM algorithm based on the rate of d.

In addition to the user profile, we estimated profiles representing the group of users for the group-based contextual suggestion and in this case, the set of documents as the input of SWLM contains the textual profile of users in the group. In this paper we take the age of users as the grouping criterion, since it has been shown as one of the most effective grouping approaches on the contextual suggestion task [2]. We discretize the age of users considering 5 years as the bin size to divide them into different groups.

3. EXPERIMENTS

In this section, we explain about our experiments and describe the results.

3.1 User Profiling Using SWLM

In the first set of experiments, we address our first research question: "How can SWLM help to estimate better user profiles for the contextual suggestion?"

After estimating user profiles using SWLM, we use the similarity of the standard language model of the candidate attractions with the estimated user profiles to score candidates and rank them. We use the inverse of JS-Divergence as the similarity function in our experiments. As the baseline, we use standard language model for user profiling in the same setting.

Table 1 presents the performance of different systems for user profiling. $UserProfile^{+}_{SWLM}$ is based on using user profiles that are estimated by significant words language model and just taking the positive user profile into account for candidate scoring. $UserProfile^{+/-}_{SWLM}$ is the same approach but taking both positive and negative profiles into consideration in order for scoring candidates. We use simple linear interpolation of the positive score with positive weight and the negative score with negative weight as the final score. $UserProfile^{+}_{SLM}$ and $UserProfile^{+/-}_{SLM}$ are baseline systems with the same settings explained above but using standard language model instead of SWLM for profiling.

As can be seen in the table, $UserProfile^{+/-}_{SWLM}$ is the best performing system compare to the others. However, the important point is that using standard language model as the profiling method, we lose performance when we take the negative profiles into account. This is due to the fact that usually negative instances are not focused on a specific topic and estimating a model representing all of them is more difficult compared to positive instances. In this case, using standard language model, we are no able to effectively capture the negative terms and the estimated negative profiles sometimes reflect positive aspects as well and by considering them in the system we penalize the positive candidates. However, SWLM is able to capture the essential terms reflecting the shared commonalities of negatively rated candidates and accidental positive terms in negative candidates are removed from the final estimation. Then, using these significant negative terms as negative profile, we are able to improve the performance in general.

Table 2: Performance of Contextual Suggestion Using Different Group Profiling Approaches

Method	ndcg@5	recip_rank	P@5
$UserProfile^{+/-}_{SWLM}$	0.2711	0.5453	**0.3911**
$UP + GroupProfile_{SLM}$	0.2612	0.5345	0.3791
$UP + GroupProfile_{SWLM}$	**0.2756**	**0.5501**	0.3811

3.2 Group Profiling Using SWLM

Here, we address our second research question: "In what conditions does information of group profiles estimated using SWLM improve the performance of the contextual suggestion?"

We group users based on their age (5 years as the bin size) into different categories, and then used SWLM again to estimate profile for each group (as a set of users) with regards to the estimated user profiles for individual users in the group. We estimate just positive group profiles according to the positive user profiles estimated with the settings of $UserProfile^{+}_{SWLM}$ system, explained in the previous section. Afterward, given an individual user, we take all the candidate suggestions and the profile of the group that the user belongs to, and then we generate group-based scores of suggestions using the similarity of the group profile with the candidate model employing inverse JS-Divergence. Then, having the group-based score, we combine it with the score achieved using the user profile in the settings of $UserProfile^{+/-}_{SWLM}$ system, using linear interpolation.

As the baseline for group profiling approach, we concatenate all the textual information of all members in the group into one document representing the whole group and then we use a standard language model to represent this document as a distribution over terms, and take this model as the group profile.

Table 2 presents the performance of contextual suggestion without help of group profiling ($UserProfile^{+/-}_{SWLM}$, the best performing system in the previous section), with help of group profiling, using SWLM as the group profiling approach ($UP + GroupProfile_{SWLM}$), and using standard language model as the group profiling approach ($UP + GroupProfile_{SLM}$).

As can be seen, using group profiles helps to improve the performance of contextual suggestion in case of using SWLM as the profiling approach but using standard language model as the profiling approach we are not able to improve the performance. However, the improvement we get from employing group profiles estimated by SWLM is not significant. We looked into the data to see in which cases adding group information helps and in which cases it is not effective. We observed that there is a correlation between the amount of improvement in contextual suggestion using group information and the rating behavior of users.

To do so, we simply take the average rate that user gave to different attractions as their general tendency of rating. Figure 1 shows the scatter plot of the change in $p@10$ after employing group-based information based on different rating tendency. According to the plot, group-based information works better when the user has a neural tendency in her rating (around rate 2) and it is less likely to help when users have rather strong biases by rating attraction with high or low rates. This could be due to the fact that in case of having neutral user, we have less string information coming from his/her profile and then group-based information is compensating this lack of strong signals

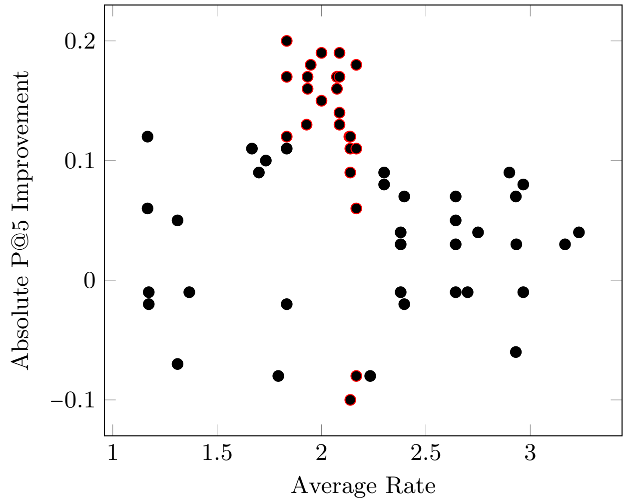

Figure 1: Improvement of the performance of contextual suggestion by help of group profiles for users with different rating behavior.

4. CONCLUSIONS

In this paper, we presented the participation of University of Amsterdam, ExPoSe team, in the TREC 2016 Contextual Suggestion Track. We described our approach which is employing Significant Words Language Models (SWLM) [1] as an effective method for estimating models representing significant features of sets of attractions as user profiles and sets of users as group profile.

We had two main research questions. The first research question was "How can SWLM help to estimate better user profiles for the contextual suggestion?" We observed that using SWLM, we are able to better estimate a model representing the set of preferences positively rated by users as their profile, compared to the case we use standard language model as the profiling approach. We also found that using negatively rated attractions as negative samples along with positively rated attractions as positive samples, we may loose the performance when we use standard language model as the profiling approach. While, using SWLM, taking negatively rated attractions into consideration may help improving the quality of suggestions.

Our second research question was "In what conditions does information of group profiles estimated using SWLM improve the performance of the contextual suggestion?" We investigated the effect of employing information from groups that users belong to on the performance of suggestions provided for individual users. We noticed that group-based suggestion helps more when users have tendency to rate attraction in a neutral way, compared to the case users are subjective in their rating behaviour.

5. ACKNOWLEDGMENTS

This research is funded in part by Netherlands Organization for Scientific Research through the *Exploratory Political Search* project (ExPoSe, NWO CI # 314.99.108), and by the Digging into Data Challenge through the *Digging Into Linked Parliamentary Data* project (DiLiPaD, NWO Digging into Data # 600.006.014).

References

[1] M. Dehghani, H. Azarbonyad, J. Kamps, D. Hiemstra, and M. Marx. Luhn revisited: significant words language models. In *The Proceedings of The ACM International Conference on Information and Knowledge Management (CIKM'16)*, 2016.

[2] M. Dehghani, H. Azarbonyad, J. Kamps, and M. Marx. Generalized group profiling for content customization. In *CHIIR'16*, pages 245–248, 2016.

[3] S. H. Hashemi, M. Dehghani, and J. Kamps. Parsimonious user and group profiling in venue recommendation. In *TREC 2015*. NIST, 2015.