Neural Endorsement Based Contextual Suggestion

Seyyed Hadi Hashemi1  Nawal Ould Amer2  Jaap Kamps1

1University of Amsterdam, Amsterdam, The Netherlands, \{hashemi|kamps\}@uva.nl
2University of Grenoble, Grenoble, France, nawal.ould-amer@imag.fr

ABSTRACT

This paper presents the University of Amsterdam’s participation in the TREC 2016 Contextual Suggestion Track. In this research, we have studied a personalized neural document language modeling and a neural category preference modeling for contextual suggestion using available endorsements in TREC 2016 contextual suggestion track phase 2 requests. Specifically, our main aim is to answer the questions: How to model users’ profiles by using the suggestions’ endorsements as an additional data? How effective is using word embeddings to boost terms’ weights relevant to the given endorsements? How to model users’ attraction-category preferences? How effective is using deep neural networks to learn users’ category preferences in contextual suggestion task? Our main findings are the following: First, the neural personalized document based user profiling using word embeddings improves the baseline content-based filtering approach based on all the common IR measures including TREC 2016 Contextual Suggestion official metric (NDCG@5). Second, neural users’ category preference modeling beats both baseline content-based filtering and the user profiling model using word-embeddings in terms of all the common IR measures.

Moreover, TREC 2016 contextual suggestion track organizers released related tags of each attraction in the qrels file. These endorsements in a way classify suggestions, which is potentially good source of information to improve users’ preference modeling. In this paper, we mainly focus on how to build and use the tag preference model in order to build contextual suggestion systems.

In TREC 2016, contextual suggestion track organizers distributed the TREC contextual suggestion web corpus, which is a web archive of the released TREC Contextual Suggestion data collection being used in both TREC 2015 and 2016 [2, 9]. In this study, we have indexed and used the released corpus as a dataset in our experiments.

TREC 2016 contextual suggestion allowed participants to participate in the contextual suggestion phase 1 or phase 2 experiments. In the phase 1 experiment, participants return a list of attraction IDs from the TREC 2016 contextual suggestion collection, but in the phase 2 experiment, participants were allowed to use the open web to retrieve suggestion candidates. All of them used the webpages of the aggregator websites such as Yelp, Google Places, Foursquare and TripAdvisor. A considerable fraction of the participants used category of suggestion candidates that is available in the Yelp website. In that track, the given context had geographical and temporal aspects.

In the TREC 2013 and TREC 2014, the participants could use either the open web or the ClueWeb12 dataset, but there were only seven submitted runs out of 34 in 2013 and 6

1. INTRODUCTION

In this paper, we present the University of Amsterdam participation in the TREC 2016 Contextual Suggestion Track [9]. The main goal of this track is to investigate search techniques for complex information needs that are highly dependent on context and user interests. In each run, participants have to produce a ranked list of suggestions for each pair of profile and context.

Each profile corresponds to a user who has judged suggestions given in a specific context. The user profiles contain a five-point scale rating for each pair of profile and example suggestions. The context provided in TREC 2016 is similar to the context being used in the TREC 2015 Contextual Suggestion Track, and it consists of a city name which represents which city the trip will occur in and several pieces of data about the trip.

In particular, in TREC 2016, the contextual suggestion organizers provide a city the trip will occur in, a trip type, a trip duration, a type of group the person is travelling with, and the season the trip will occur in as contexts of the venue recommendation task. Hopefully, almost all of the given contextual suggestion requests have information about all types of the mentioned contexts, which makes it a very interesting data to test contextual suggestion systems.

In TREC 2016, contextual suggestion track organizers distributed the TREC contextual suggestion web corpus, which is a web archive of the released TREC Contextual Suggestion data collection being used in both TREC 2015 and 2016 [2, 9]. In this study, we have indexed and used the released corpus as a dataset in our experiments.

In this paper, our main aim is to study the following research questions:

1. How to effectively model users’ profiles using neural language modeling?
2. How to effectively learn users’ preferences using neural category preference modeling?

The rest of this paper is organized as follows. In Section 2 we review some related work on Contextual Suggestion track. In Section 3 we detail our models of Contextual Suggestion, and Section 4 is devoted to the TREC contextual suggestion results. Finally, we present the conclusions in Section 5.

2. RELATED WORK

In the TREC 2012 Contextual Suggestion Track, participants were allowed to use the open web to retrieve suggestion candidates. All of them used the webpages of the aggregator websites such as Yelp, Google Places, Foursquare and TripAdvisor. A considerable fraction of the participants used category of suggestion candidates that is available in the Yelp website. In that track, the given context had geographical and temporal aspects.

In the TREC 2013 and TREC 2014, the participants could use either the open web or the ClueWeb12 dataset, but there were only seven submitted runs out of 34 in 2013 and 6
Algorithm 1: Estimating a Personalized Document Language Model

1: procedure PDL(d, TG_u(d))
2: for each t ∈ d do
3:     P(t|θ_u) = \frac{tf(t,d)}{|d|} \sum_{t \in TG_u(d)} P(t|θ_u)P(t|tg)
4: end for
6: end procedure

Algorithm 2: Estimating a User Model

Require:
- Du = \{d_1, d_2, ..., d_n\} Document preference of user u.
- Vu = \{t_1, t_2, ..., t_M\} User vocabulary.

Ensure:
- θ_u User Model.

1: for each d ∈ Du do
2:     θ_d_u ← PDL(d, TG_u(d))
3: end for
4: for each t ∈ Vu do
5:     for each d ∈ Du do
6:         P(t|θ_u) = \frac{1}{|Du|} \sum_{d \in Du} P(t|θ_u)P(t|tg)
7: end for
8: end for

where tf(t,d) is a frequency of term t in the document d, and |d| is a document length.

Then, we estimate the personalized document model θ_d_u using tags TG_u(d) as follow:

\[ P(t|θ_u) = \frac{1}{|TG_u(d)|} \sum_{tg \in TG_u(d)} P(t|θ_u)P(t|tg) \]  (2)

where P(t|θ_u) is the probability of term in a document as described in Eq.1, P(t|tg) is a probability of selecting a term t given a tag tg, and |TG_u(d)| number of tag assigned to the document d by a user u.

The probability P(t|tg) is computed using the cosine similarities between the two embedded vectors corresponding to term t and tag tg as follow:

\[ P(t|tg) = sim(t, tg) \]  (3)

3.1 User Profiling Using Word Embeddings

This section studies how to effectively model users’ profiles to be used in the content based filtering systems, aiming to answer our first research question: How to effectively model users’ profiles using neural language modeling?

3.1.1 Personalized Document Language Model

In this part, we present how to personalize a document model using user tags. The goal is to estimate a best personalized term distribution for the document according to the tags assigned by a user.

Our approach is shown in Algorithm 1. Given a document d = \{t_1, t_2, ..., t_n\} and his related tags TG_u(d) = \{tg_1, tg_2, ..., tg_m\} assigned by a user u. We first estimate a document model θ_d as first estimation using maximum likelihood as follow:

\[ P(t|θ_d) = \frac{tf(t,d)}{|d|} \]  (1)
To learn an effective model and avoid over-fitting, we have implemented a content-based filtering baseline using standard language model to model users’ profiles, and we have used KL-divergence to estimate relevance of the suggestion candidate to the user profile. Table 1 shows our submissions in Section 3 is discussed. These results are based on the official TREC 2016 Contextual Suggestion track qrels.

In order to evaluate our proposed models, we have implemented a content-based filtering baseline using standard language model to model users’ profiles, and we have used KL-divergence to estimate relevance of the suggestion candidate to the user profile. Table 1 shows our submissions results against the content-based filtering baseline.

As it is shown in Table 1 both the proposed neural approaches beat the baseline in terms of all the common IR measures. The category of the attractions proves to be better than the two others phase 2 runs.

Table 1: TREC 2016 Contextual Suggestion Track: Phase 2 results.

<table>
<thead>
<tr>
<th>RunID</th>
<th>NDCG@5</th>
<th>P@5</th>
<th>MRR</th>
<th>NDCG</th>
<th>MAP</th>
<th>bpref</th>
<th>P@10</th>
<th>Rprec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.1967</td>
<td>0.2862</td>
<td>0.4440</td>
<td>0.6257</td>
<td>0.3862</td>
<td>0.4332</td>
<td>0.3086</td>
<td>0.3551</td>
</tr>
<tr>
<td>UAmsterdamCB</td>
<td>0.2730</td>
<td>0.4069</td>
<td>0.5631</td>
<td>0.6499</td>
<td>0.4076</td>
<td>0.4337</td>
<td>0.4000</td>
<td>0.3780</td>
</tr>
<tr>
<td>UAmsterdamDL</td>
<td>0.2824</td>
<td>0.4448</td>
<td>0.5924</td>
<td>0.6544</td>
<td>0.4168</td>
<td>0.4452</td>
<td>0.4310</td>
<td>0.3881</td>
</tr>
</tbody>
</table>

For the classification purpose and having probabilities as outputs, we have used Logistic classifier in the last layer. We use variable \( c \in \{0, 1\} \) to show relevance of a suggestion to the given user in a context. Specifically, \( P_b(c = 1|u, c, s) \) is the relevance score of the suggestion \( s \) to the user \( u \) and context \( c \), in which \( \theta \) is unknown parameters that are learned using maximum likelihood estimation (MLE) based on the train set, which consists of the users’ preferences available in the profile of each request.

Given the relevance judgments \( r \) of each suggestion \( s_k \) to a user \( u_i \) and context \( c_j \) in the users’ profiles available at each requests, the likelihood \( L \) of the train set is as follows:

\[
L = \prod_{i=1}^{U} \prod_{|C|} \prod_{|S|} P_{\theta}(c = 1|u_i, c_j, s_k) \prod_{P_{\theta}(c = 0|u_i, c_j, s_k)}^{1-r},
\]

in which we assume relevance judgments \( r \) are generated independently. We model \( P_{\theta}(c = 1|u_i, c_j, s_k) \) by logistic function on a linear combination of inputs from the last hidden layer units’ outputs. Then, the unknown parameters \( \theta \) are optimized by maximizing the following log likelihood function:

\[
\theta^* = \arg\max_\theta \sum_{i=1}^{U} \sum_{|C|} \sum_{|S|} r \log P_{\theta}(c = 1|u_i, c_j, s_k) + (1 - r) \log P_{\theta}(c = 0|u_i, c_j, s_k).
\]

4. RESULTS

In this section, the result of the two approaches detailed in Section 3 is discussed. These results are based on the official TREC 2016 Contextual Suggestion track qrels.

As it is shown in Table 1 both the proposed neural approaches beat the baseline in terms of all the common IR measures. The category of the attractions proves to be very useful to include in the contextual suggestion systems, explaining why the “UAmsterdamDL” approach performed better than the two others phase 2 runs.

5. CONCLUSION

In this paper, we studied contextual suggestion problem through neural user profiling and neural category preference modeling by the help of suggestions’ endorsements. According to the phase 2 results of the TREC 2016 contextual suggestion track, using word embeddings to boost terms’ weights related to suggestions’ endorsements improves baseline content-based filtering approach in the contextual suggestion problem based on all common IR measures. Moreover, phase 2 results show that neural category preference...
modeling of the users can lead to even better results than the other tested user modeling approaches in contextual suggestion task. Specifically, the contextual suggestion submitted run based on neural category preference modeling performs better than our user profiling based submission and the content-based filtering baseline in terms of all the common IR measures including the TREC contextual suggestion official evaluation metric (NDCG@5). As a future work, we continue to work on users’ preference modeling using category profiles created based on word embeddings.

Acknowledgments
This research is funded in part by the European Community’s FP7 (project meSch, grant # 600851), the Netherlands Organization for Scientific Research (ExPoSe project, NWO CI # 314.99.108; DiLiPaD project, NWO Digging into Data # 600.006.014), French Région Auvergne Rhône-Alpes ReSPIr project, and ELIAS (ESF Research Networking Programme).

REFERENCES