WHUIRGroup at the CLEF 2016 eHealth Lab Task 3

Ruixue Wang, Wei Lu, Ke Ren

School of Information Management Wuhan University,
No. 299 Bayi Road, 430072 Wuhan, Hubei, China
{ruixue_wang , weilu}@whu.edu.cn renke19910123@163.com

Abstract. This paper presents our work on the 2016 CLEF eHealth Task 3.We used Indri to conduct our experiments. We used CHV to expand query and proposed a learning-to-rank algorithm to re-rank the result.
Keywords: CHV-based query expansion, learning-to-rank , language model

1 Introduction

More and more people through search engine find medical information to diagnose their symptoms and understand the health information. Such searchers we call laypeople or health consumer have little medical knowledge and often fail to find the necessary information. Because they are unfamiliar with medical terminology and the reliability the web sources, when laypeople try to use the results for self-treatment it may lead to dangersous consequences.

To help laypeople, Conference and Labs of the Evaluation Forum (CLEF) launched the eHealth Evaluation Lab in 2013[1].The 2016 CLEF eHealth Task 3 Patient-Centred Information Retrieval[2][3] is a continuation of previous CLEF eHealth IR tasks. In this year's task ,the dataset used is ClueWeb12 B13 instead of previous data collection which is one million documents provided by the Khresmoi project[1].The queries are generated by query generators who read the real health consumer posts from health web forums "askDocs[2]" .

In this paper we present our participation to the 2016 CLEF eHealth Task 3. The rest of this paper is organized as follows. Section 2 presents the method we used to solve the task. The experimental results are described in Section 3.We conclude in Section 4.

[1] Medical Information Analysis and Retrieval, http://www.khresmoi.eu
[2] https://www.reddit.com/r/AskDocs/

2 Method

2.1 Framework of our system

Figure 1 presents our framework of integration of readability of the results and the learning-to-rank algorithm. We use the CLEF 2015 task 2 readability relevance data to train readability model and use the model to predict the readability of each running results. The score of readability is one feature in the learning-to-rank model. We also use the weighting score and the rank of each document-query pair from the language model as the features to train the learning-to-rank model. We use the CLEF 2015 task 2 running results as the training data.

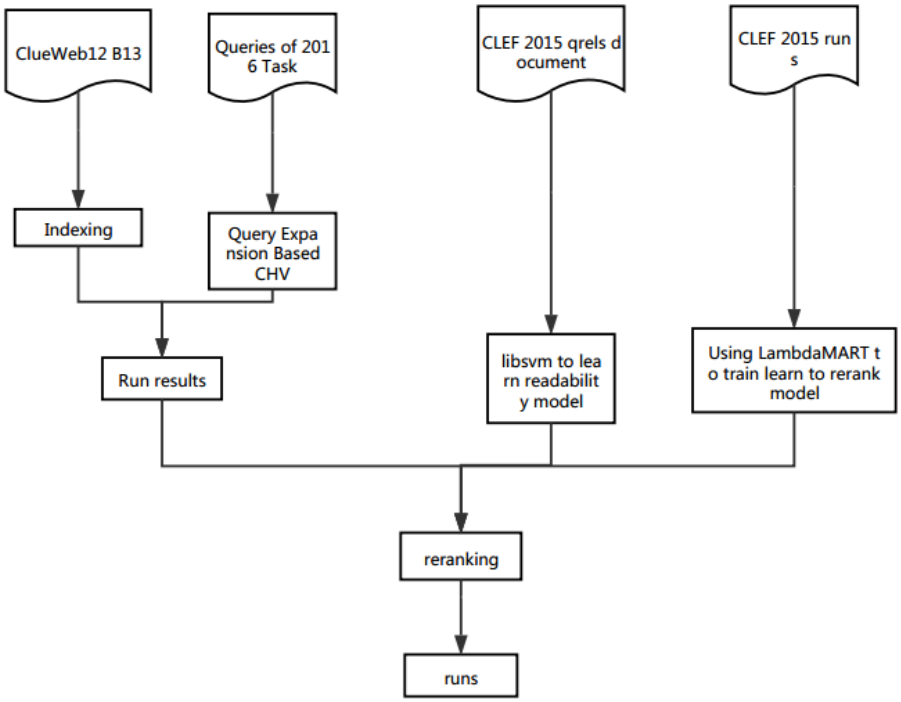

Fig. 1. Framework

2.2 CHV-based query expansion

This year, the organizer extracted posts from the 'askDocs' forum of Reddit, and presented them to query generators who had to create queries based on what they read in the initial user post. So the queries are close to laypeople's queries in search engine.

But different people may use different terms to express their symptoms, so we adopt Consumer Health Vocabularies (CHV)[3] to expand queries.

Consumer Health Vocabularies are developed by Qing T. Zeng [4] .It contains different ways laypeople express and think about health topics and bridges this vocabularies to UMLs expressions. We used each term in queries and search them in online CHV systems and the term expressions which is different from the initial term are used for query expansion.

2.3 Readability of the search result

In CLEF 2015 eHealth IR task, readability was an important factor to judge the retrieval results and used Urbp to evaluate the result. Readability assessments were collected by asking the assessors whether they believed a patient would understand the retrieved document. Assessments were provided on a four point scale, the point meaning showed in Table1.

Table 1.

Point	Meaning
1	It is very technical and difficult to read and understand
2	It is somewhat technical and difficult to read and understand
3	It is somewhat easy to read and understand
4	It is very easy to read and understand

Tiffany M Walsh[5] used SMOG, Gunning FOG and Flesch-Kincaid to compare the readability of Internet-based consumer health information articles from the associations that represent the 5 leading casue of health-related death in America. So we take the scores of SMOG, Gunning FOG, Flesch-Kincaid and the number of words, complex words, sentences, characters and syllable, as the features to train the model of readability. We remove the html tag of the document and use the same way to remove tags of the results of CLEF2016 task 3 runs. We used libsvm[4] tools to train the model.

2.4 Learning-to-rank Algorithm

We used the 2015 best run as training data and RankLib [5] to train the learning-to-rank model. We adopted LambdaMART[6] method and use readability features and LM result ranking features to train the model. Then apply the model to classify the result of CHV-based query expansion result. Finally, the results are re-ranked by their new scores.

[3] http://consumerhealthvocab.chpc.utah.edu/CHVwiki/
[4] https://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html
[5] https://sourceforge.net/p/lemur/wiki/RankLib/

3 Experiments and results

The dataset of 2016 eHealth Task 3 is ClueWeb12 B13. The organizer provide access to an Azure instance where participants can access the dataset of ClueWeb12 B13, standard indexes built with the Terrier tool and the Indri tool and additional resources such as a spam list, anchor texts, urls. We adopt Indri Index in the Azure to conduct our experiments. We submit three runs where the description for each run is in Table2.

Table 2.

Run id	description
1	The baseline with the Language Model
2	Using the CHV to obtain the expanded query ,then search with Language Model
3	Utilizing the learning-to-rank model to re-rank the RUN2

4 Conclusions

In CLEF 2016 eHealth Lab, our group focus on the task 3 which is patient-centred information retrieval. We used language model as the baseline.Because of the gap between laypeople and professions, we propose a CHV-based query expansion model. When we train the re-rank model, we consider readability as one of the features to make laypeople easily read and understand the result. In the future, we will continue the CHV-based query expansion for better retrieval results.

References

1. Goeuriot, L., Jones, G.J.F., Kelly, L., Leveling, J., Hanbury, A., M ̈uller, H., Salantera, S., Suominen, H., Zuccon, G.: ShARe/CLEF eHealth Evaluation Lab 2013, Task 3: Information retrieval to address patients' questions when reading clinical reports. In: CLEF 2013 Online Working Notes. (2013)
2. Kelly, Liadh and Goeuriot, Lorraine and Suominen, Hanna and Névéol, Aurélie and Palotti, Joao and Zuccon, Guido. Overview of the CLEF eHealth Evaluation Lab 2016. CLEF 2016 - 7th Conference and Labs of the Evaluation Forum, Lecture Notes in Computer Science (LNCS), Springer, September, 2016.
3. Zuccon, Guido and Palotti, Joao and Goeuriot, Lorraine and Kelly, Liadh and Lupu, Mihai and Pecina, Pavel and Mueller, Henning and Budaher, Julie and Deacon, Anthony. The IR Task at the CLEF eHealth Evaluation Lab 2016: User-centred Health Information Retrieval. CLEF 2016 Evaluation Labs and Workshop: Online Working Notes, CEUR-WS, September, 2016.
4. Zeng Q T, Tse T. Exploring and developing consumer health vocabularies[J]. Journal of the American Medical Informatics Association, 2006, 13(1): 24-29.
5. Walsh T M, Volsko T A. Readability assessment of internet-based consumer health information[J]. Respiratory care, 2008, 53(10): 1310-1315.

6. Q. Wu, C.J.C. Burges, K. Svore and J. Gao. Adapting Boosting for Information Retrieval Measures. Journal of Information Retrieval, 2007.