Team DA_IICT at Consumer Health Information Search @FIRE2016

Jainisha Sankhavara
IRLP Lab
Dhirubhai Ambani Institute of Information and Communication Technology
Gandhinagar, Gujarat, India
jainishasankhavara@gmail.com

ABSTRACT

Consumer Health Information Search task focuses on retrieval of relevant multiple perspectives for complex health search queries. This task addresses the queries which do not have a single definitive answer but having diverse point of views available. This paper reports the result of standard retrieval methods for identifying the aspects of retrieval result towards the query.

Keywords

Consumer Health Information Search, Health Information Retrieval

1. INTRODUCTION

People are highly using web search engines for health information retrieval now a days. These search engines are quite suitable to answer the straightforward health related medical queries but some queries are complex in a way that they do not have a single definitive answer, instead they have multiple perspectives to the queries, both for and against hypothesis. The presence of multiple perspectives with different grades of supporting evidence (which is dynamically changing over time due to the arrival of new research and practice evidence) makes it all the more challenging for a lay searcher. Consumer Health Information Search (CHIS) aims to target such information retrieval search tasks, for which there is no single best correct answer but having multiple and diverse perspectives/points of view available on the web regarding the queried information.

The description of data is provided in section 2. The experiments and results are described in section 3 and section 4 respectively and we conclude in section 5.

2. CHIS TASK

There will be two sets of tasks:

A) Given a CHIS query, and a document/set of documents associated with that query, the task is to classify the sentences in the document as relevant to the query or not. The relevant sentences are those from that document, which are useful in providing the answer to the query.

B) These relevant sentences need to be further classified as supporting the claim made in the query, or opposing the claim made in the query.

Example query: Are e-cigarettes safer than normal cigarettes?

S1: Because some research has suggested that the levels of most toxicants in vapor are lower than the levels in smoke, e-cigarettes have been deemed to be safer than regular cigarettes. A)Relevant, B) Support

S2: David Peyton, a chemistry professor at Portland State University who helped conduct the research, says that the type of formaldehyde generated by e-cigarettes could increase the likelihood it would get deposited in the lung, leading to lung cancer. A)Relevant, B) oppose

S3: Harvey Simon, MD, Harvard Health Editor, expressed concern that the nicotine amounts in e-cigarettes can vary significantly. A)Irrelevant, B) Neutral

There were 5 queries provided and 357 sentences across those queries. The performance is measured in terms of percentage accuracy of each task against each query and a task wise average over all five queries are considered as evaluation measure.

3. EXPERIMENTS

The experiments include standard retrieval methods to identify relevant and irrelevant sentences (task A). To identify weather the sentence is supporting the claim or opposing the claim (task B), the standard query expansion technique was used.

The experiments are done using terrier[2] tool-kit which are openly available. The experiments focuses on how useful the standard retrieval methods are to identify the relevance at sentence level instead of documents and how it can be useful to identify the supporting or opposing nature of sentences to the hypothesis of query. BM25[1],[3] model is used to identify relevant/not-relevant sentences and TF-IDF[1] with query expansion is used to identify supporting/opposing nature of the sentences.

Task A: Identify relevant/non-relevant sentences.

The sentences are indexed using terrier and the retrieval is performed against each queries using BM25 retrieval model.

The retrieved sentences are marked as relevant for task A and others (non-retrieved sentences) are considered as non-relevant to the query.

Task B: Identify support/oppose/neutral nature of sentences.

The sentences are indexed using terrier and the queries are executed against indexed sentences using TF-IDF retrieval model with Bo1 query expansion model taking top 5 sentences as feedback and 30 terms as expansion terms. The sentences retrieved using query expansion are which are identified relevant according to task A are marked as supporting and the sentences which are not retrieved using query expansion but retrieved using task A are marked as opposing to the query since they are relevant to the query. All irrelevant sentences are considered to be neutral to the query.

4. RESULTS

The percentage accuracy of the query wise results obtained by above described method is given in the below table.

Query	Task A	Task B
Skincare	52.27272727	37.5
MMr	87.93103448	46.55172414
HRT	91.66666667	27.77777778
Ecig	54.6875	46.875
Vitc	64.86486486	31.08108108
Overall	70.28455866	37.9571166

Table 1: Percentage accuracy for both the tasks

There were 9 teams participated in task A and 8 teams in task B. The comparison of the overall percentage accuracy of the results with maximum of all teams and average of all teams is given in the following graph.

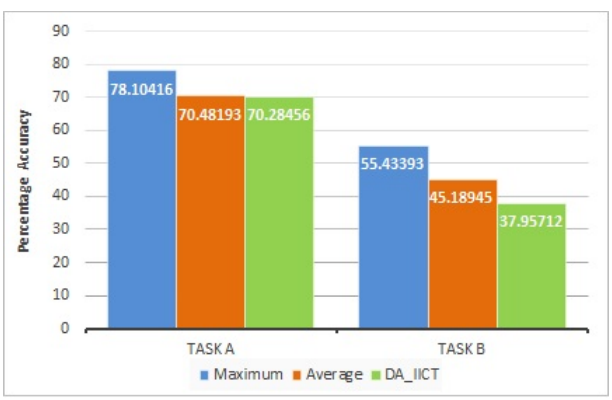

Figure 1: Comparison of with maximum and average results

The results of task A are comparable to the average of all other systems that means standard information retrieval algorithms are helpful to get average results but for task B, standard information retrieval algorithms fails to achieve atleast average results. So, the standard algorithms are less recommendable to use to extract supporting/opposing sentences but definitely can be used to extract relevant/non-relevant sentences.

5. CONCLUSION

The paper describes results of standard information retrieval algorithms on complex medical queries which have multiple perspectives available. Standard information retrieval algorithm gives average results when identifying relevant/ non-relevant sentences but it gives less than the average results in identifying supporting/opposing sentences. In task A, our results are third in the rank-list of all participants.

6. REFERENCES

[1] I. Mogotsi. Christopher d. manning, prabhakar raghavan, and hinrich schütze: Introduction to information retrieval. *Information Retrieval*, 13(2):192–195, 2010.

[2] I. Ounis, G. Amati, V. Plachouras, B. He, C. Macdonald, and D. Johnson. Terrier information retrieval platform. In *European Conference on Information Retrieval*, pages 517–519. Springer, 2005.

[3] S. Robertson and H. Zaragoza. *The probabilistic relevance framework: BM25 and beyond*. Now Publishers Inc, 2009.