Exploiting Term Dependence while Handling Negation in Medical Search

Nut Limsopatham[1], Craig Macdonald[2], Richard McCreadie[2], Iadh Ounis[2]
nutli@dcs.gla.ac.uk[1], firstname.lastname@glasgow.ac.uk[2]
School of Computing Science
University of Glasgow, Glasgow, UK

ABSTRACT

In medical records, negative qualifiers, e.g. no or without, are commonly used by health practitioners to identify the absence of a medical condition. Without considering whether the term occurs in a negative or positive context, the sole presence of a query term in a medical record is insufficient to imply that the record is relevant to the query. In this paper, we show how to effectively handle such negation within a medical records information retrieval system. In particular, we propose a term representation that tackles negated language in medical records, which is further extended by considering the dependence of negated query terms. We evaluate our negation handling technique within the search task provided by the TREC Medical Records 2011 track. Our results, which show a significant improvement upon a system that does not consider negated context within records, attest the importance of handling negation.

Categories and Subject Descriptors: H.3.3 [Information Storage & Retrieval]: Information Search & Retrieval

General Terms: Experimentation, Performance

Keywords: Medical Retrieval, Negation, Term Dependence

1. INTRODUCTION

Search in the medical domain is notable for its extensive use of negated language within medical records [2]. However, recently Koopman et al. [2] asserted that the use of term frequency in document weighting models such as BM25 automatically alleviates the problems of negated language. Notably, they argue that terms in negative contexts typically appear once per medical record while, in contrast, terms indicating the presence of a medical condition (positive terms) appear frequently. In this paper, we argue that a dedicated negation handling approach will outperform traditional retrieval models for medical domain queries. Additionally, prior works have focused almost exclusively on negated language in queries rather than documents (e.g. boolean retrieval models or vector negation [7]). For example, for the query 'NOT chest pain AND shortness of breath', the boolean model will find documents not containing 'chest pain', thereby not considering any negation occurring within the records. Moreover, most words indicating negation (e.g. no, not) are stopwords, which are not typically indexed [1].

To cope with negation in medical search, we propose a two-step process: a term representation approach, *NegFlag*, to facilitate the handling of negative context in medical records; and a novel term dependence approach to demote medical records containing query terms in a negated context. We evaluate our proposed approach using the TREC 2011 Medical Records track [6] test collection. As our results exhibit significant improvement over a baseline where negation is not handled, we conclude that negated language should be explicitly handled for effective medial records search.

SIGIR'12, August 12–16, 2012, Portland, Oregon, USA.
ACM 978-1-4503-1472-5/12/08.

Original record	Patient reports palpitations but does not have fever
Negation detection	Patient reports palpitations but does *not have fever*
Removing stopwords	Patient reports palpitations *fever*
NegFlag representation	Patient reports palpitations *n0fever*

Table 1: The NegFlag process for a medical record - italicised terms occur in a negative context.

2. THE NEGFLAG APPROACH

Our negated term representation approach, *NegFlag*, modifies the indexing process to distinguish between positive and negative context terms in medical records, which are identified using the NegEx algorithm [1]. Identified negated terms are replaced with special negated versions of those terms. Table 1 shows how an example sentence is processed using NegFlag, such that the term 'fever' is replaced with its negated version, 'n0fever'.

Hence, given the query 'find patient with fever', without NegFlag, a retrieval model might erroneously rank the original record in Table 1 highly, because it contains all of the query terms. However, after NegFlag processing, as 'fever' has become 'n0fever', the record would not score as highly.

However, while NegFlag only helps to retrieve records containing query terms with positive contexts, it does not prevent records with the negated occurrences from being retrieved. For example, the NegFlag processed example in Table 1 might still be retrieved for the query 'find patient with fever and palpitations', when the patient is known not to have a fever. To alleviate this, in the next section, we propose the use of term dependence to demote medical records containing query terms within negative contexts.

3. TERM DEPENDENCE FOR NEGATION

Term dependence (e.g. Markov Random Fields [4]) has been used to improve effectiveness by scoring higher documents containing many occurrences of pairs of query terms in close proximity. In contrast, we propose to use term dependence to *demote* records containing the *negated* form of neighbouring terms occurring in the queries. For example, given a query 'chest pain', documents containing the pair of terms 'n0chest n0pain' should be demoted. To this effect, we score medical record r for a query Q, taking negation into account, as follows:

$$score\,(r,Q) = \sum_{t \in Q} score(r,t) - \sum_{\langle t_1,t_2\rangle \in Q'} score(r, \langle t_1, t_2\rangle) \quad (1)$$

There are two components in Equation (1), namely the positive scoring of positive query terms, and the negative term dependence score for the negated forms of the query terms. $score(r,t)$ is the score assigned to a query term t in medical record r using any term weighting model, Q' is the set of negated forms of the positive query terms in Q, and $\langle t_1, t_2\rangle$ is a pair of negated terms in Q'. Two types of term dependence are possible [4, 5]: for full dependence (FD), $\langle t_1, t_2\rangle$ is the set that contains unordered pairs of neighbouring terms; for sequential dependence (SD), $\langle t_1, t_2\rangle$ is the set that contains ordered pairs of neighbouring terms. For $score(r, \langle t_1, t_2\rangle)$, we use the binomial randomness model pBiL [5] from the Divergence from Randomness (DFR) framework to score the occurrences of a pair of terms within *window_size* tokens in a medical record r.

4. EXPERIMENTAL RESULTS

We evaluate our negated term representation and the term dependence approaches using the 34 topics from the TREC 2011 Medical Records track [6]. In this track, the task is to identify relevant patient *visits* for each topic, where a visit contains all of the medical records associated with that patients' admission to the hospital. For indexing and retrieval, we use Terrier[1], applying Porter's English stemmer and removing stopwords. The parameter-free DFR DPH term weighting model is used to rank medical records. The expCombSUM voting technique [3] is then used to rank visits based on the scores of their associated medical records. The number of voting medical records is limited to 5,000 as this was found to be effective in preliminary experiments. We hypothesise that negation should be explicitly handled, hence we compare our approach with a baseline where negation is not explicitly handled, as suggested in [2].

Figure 1 shows the bpref retrieval performance comparing NegFlag, as well as the SD and FD variants of term dependence with NegFlag, and the baseline where negation is not handled. From Figure 1, we observe that our approach employing either only NegFlag or both NegFlag and term dependence outperforms the baseline (bpref 0.4871), by up to 4%. Indeed, NegFlag alone markedly improves the retrieval performance over the baseline, while term dependence for *window_size* = 3 using either SD or FD results in statistically significant improvements over the baseline (paired t-test, $p < 0.05$).

For SD, small *window_size*s are more effective, but performance is generally stable across different *window_size*s, suggesting that the presence of negated ordered pairs anywhere in a medical record is sufficient to ascertain if it should be demoted. For FD, *window_size* > 3 degrades performance compared to NegFlag, but still outperforms the baseline.

Next, we further evaluate the effectiveness of our approach after applying a query expansion (QE) technique (namely DFR Bo1 from Terrier). Table 2 shows the retrieval performances (in terms of bpref, precision@10, and R-precision) of NegFlag, with SD and FD term dependence *window_size* = 3 (identified best settings in Figure 1), as well as without term dependence, compared to a baseline applying only QE. The performances of the top 3 best systems at TREC 2011 are also reported. We observe that the QE baseline is outperformed by NegFlag for all the measures, however,

[1] `http://terrier.org`

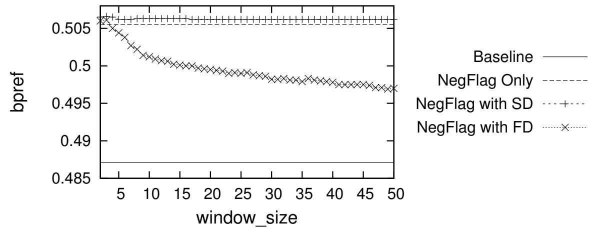

Figure 1: bpref performances of the baseline and our approach, while varying *window_size*.

Approach	QE	bpref	P@10	R-prec
Baseline	Bo1	0.5264	0.6147	0.4290
NegFlag	Bo1	**0.5436**	**0.6324**	**0.4351**
NegFlag with SD *window_size* = 3	Bo1	0.5420	**0.6324**	0.4337
NegFlag with FD *window_size* = 3	Bo1	0.5433	0.6235	0.4332
CengageM11R3	N/A	**0.5520**	**0.6560**	**0.4400**
SCAIMED7	N/A	0.5520	0.6030	0.4250
UTDHLTCIR	N/A	0.5450	0.6030	0.4220

Table 2: Performances of NegFlag and negative term dependence with QE, and the top 3 best systems from TREC 2011.

the combination of term dependence and query expansion when using NegFlag remains challenging. In addition, comparing with the best systems reported at TREC 2011, we find that NegFlag performs better than the second ranked group in terms of precision@10 and R-precision, while for bpref, it is comparable with the third ranked group. Importantly, this is despite our approach not requiring any of the domain-specific ontologies that are exploited in those systems. Overall, our results are very promising, particularly if term dependence and query expansion can be successfully combined in future work.

5. CONCLUSIONS

We have proposed a novel approach to handle negation in medical search using our NegFlag term representation and a novel use of term dependence to demote documents containing the query terms in a negative context. Our approach is shown to be effective on the Medical Records track test collection, across a range of window sizes. Moreover, it verifies our stance that negation should be explicitly handled in medical search. Our proposed approach could also work with queries that contain negative contexts; however, as the available topics in the TREC 2011 test collection do not contain any negated terms, we leave this evaluation for future work.

6. REFERENCES

[1] W. Chapman, W. Bridewell, P. Hanbury, G. Cooper, B. Buchanan. A simple algorithm for identifying negated findings and diseases in discharge summaries. *J. of Biomedical Informatics*, 5:301–310, 2001.

[2] B. Koopman, P. Bruza, L. Sitbon, M. Lawley. Analysis of the effect of negation on information retrieval of medical data. In *ADCS*'10.

[3] C. Macdonald, I. Ounis. Voting for candidates: adapting data fusion techniques for an expert search task. In *CIKM*'06

[4] D. Metzler, W. B. Croft. A Markov random field model for term dependencies. In *SIGIR*'05.

[5] J. Peng, C. Macdonald, B. He, V. Plachouras, I. Ounis. Incorporating term dependency in the DFR framework. In *SIGIR*'07.

[6] E. Voorhees, R. Tong. Overview of the TREC 2011 Medical Records Track. In *TREC*'11.

[7] D. Widdows. Orthogonal negation in vector spaces for modelling word-meanings and document retrieval. In *ACL*'03.