Medical opinion lexicon: an incremental model for mining health reviews

Article · January 2014

DOI: 10.7813/2075-4124.2014/6-1/A.39

7 authors, including:

Dr. Muhammad Zubair Asghar
Gomal University

Aurangzeb Khan
University of Science & Technology Bannu

Rahman Ullah

Muhammad Z. Asghar, A. Khan, Fazal M. Kundi, M. Qasim, F. Khan, R. Ullah, Irfan U. Nawaz. Medical opinion lexicon: an incremental model for mining health reviews. International Journal of Academic Research Part A; 2014; 6(1), 295-302.
DOI: *10.7813/2075-4124.2014/6-1/A.39*

Library of Congress Classification: T58.5-58.64, TP1-1185

MEDICAL OPINION LEXICON: AN INCREMENTAL MODEL FOR MINING HEALTH REVIEWS

Muhammad Zubair Asghar[1], Aurangzeb Khan[2], Fazal Masud Kundi[1], Maria Qasim[1], Furqan Khan[1], Rahman Ullah[1], Irfan Ullah Nawaz[1]

[1]Institute of Computing and Information Technology, Gomal University, D.I.Khan,
[2]University of Science and Technology Bannu **(PAKISTAN)**
E-mails: Zubair@gu.edu.pk**,** aurangzebb_khan@yahoo.com, paitank@yahoo.com, maria_qasim_dik@yahoo.com, furqan.saddozai@yahoo.com, rahman_lakki@yahoo.com, irfannawaz88@yahoo.com

DOI: 10.7813/2075-4124.2014/6-1/A.39

Received: 03 Sept, 2013
Accepted: 25 Jan, 2014

ABSTRACT

Opinion mining concentrates on retrieving opinions of online users about a service, product, and policy. In this paper, we propose a medical opinion lexicon for mining health reviews available on different health forums. The proposed technique is based on the incremental modal and corpus of health reviews by creating medical polarity lexicon for medical terms. In each increment, vocabulary of lexicon is enhanced systematically, polarity score with each word is attached, and finally, resulting lexicon is filtered from unnecessary words by using word sense disambiguation techniques. The comparative results show the efficiency of proposed method and it outperforms the existing approaches. The proposed approach achieves an accuracy of 82% on training corpus and 78% on testing corpus of health reviews.

Key words: Opinion Lexicon, Opinion Mining, Reviews, Health, Polarity Classification

1. INTRODUCTION

With the rapid increase in online health resources and growing interest of public in the extraction of opinions from such resources has attracted opinion miners to develop opinion mining applications for online medical text. Study conducted by (Fox, 2013) states that more than 85 % internet users have searched for online health information. Studies conducted by [Ma, Chen and Xiao, 2010; Jaloba, 2009; White, Andre and Tan, 2009; Sarasohn-Kahn, 2008] reports that social media has a positive impact on patient health, as patients discuss about their diseases and medications online with other people more frankly as compared to their family members. Most of studies to date have focused on extracting opinions from user generated reviews in non-medical domains such as movies, products, airlines and hotels, and comparatively less attention is given to mining health related opinions from patients due to non-availability of specialized opinion lexicon in this domain.

Opinion information contained in existing general purpose subjectivity lexicons (Baccianella, Esuli and Sebastiani, 2010; Stone, Dunphy, and Marshall Smith, 1966; Wilson, Wiebe and Hoffmann, 2009) is likely to be insufficient for special purpose domains such as medical applications, as such general purpose lexicons lack the content and concepts needed for processing specialized information.

Word sense disambiguation is a major issue faced by natural language processing experts in general and opinion miners in particular, because words with multiple senses often result in in-accurate results. Developing opinion lexicons for a specific domain need a limited vocabulary which can help significantly to resolve this problem. One of the previous studies (Goeuriot, Na, Min Kyaing, Khoo, Chang, Theng and Kim, 2012) on constructing special purpose domain lexicons rely on merging existing general purpose lexicons and then creating specialized lexicon from available corpus, which is time consuming and computationally not very much efficient.

One of the feasible solutions for these mentioned problems is to develop an efficient mechanism for automatic creation of dedicated opinion lexicon, which can act as an alternate small scale repository for existing manually annotated general and special purpose lexicons.

In this paper, we present an incremental method for creating special purpose medical opinion lexicon that can store medical related words and phrases along with their polarity scores. Unlike previous studies, our method takes a corpus containing patient reviews, a specialized medical dictionary, similarity cleansing module, a general purpose opinion lexicon, and term filtering strategy; which makes it attractive for opinion mining applications based on medical reviews developed so far.

In first increment each word present in the parsed review corpus is matched against the specialized medical dictionary to check whether a word/phrase in the corpus is a valid medical entry. In next increment, terms are further expanded using Web Thesaurus. The most nearest words are identified using Lin's measure and added to new lexicon; process is repeated iteratively and terminated when stopping condition is fulfilled. In third increment, we calculate polarity score for all the senses of each POS for a given word by using opinion lexicon and existing polarity formula. Word sense disambiguation is applied to isolate most relevant medical terms. Finally, semantically related words are filtered using second order PMI computation and appended to newly constructed lexicon.

This incremental special purpose lexicon creation technique has several advantages over existing similar approaches, such as opinion lexicon can be created rapidly, initial index words and input corpus help in extending the lexicon, important opinion words along with their opinion score are extracted and added automatically, and unrelated words are excluded from final lexicon.

We demonstrate that the resulting medical opinion lexicon is comparable to existing health related lexicons in terms of polarity assignment performance and efficient data coverage. We have performed experiments on cholesterol-controlling medication, but the method is applicable to all other drugs.

Rest of the paper is organized as: section 2 presents the related work, section 3 gives description of the corpus, lexicon creation is mentioned in section 4, and section 5 describes evaluation and presentation of results.

2. RELATED WORK

We are interested in developing subjectivity lexicon by mining non-expert (patients) opinions available on medical discussion forums. In this section we describe related work performed on building polarity lexicon from user reviews expressed on social media forums with emphasis on construction of medical subjectivity lexicon. There are many papers describing lexicon compilation using different approaches. Here we will describe the most relevant ones, with respect to our aim.

SentiWordNet (Baccianella, Esuli and Sebastiani, 2010) is a popular opinion mining polarity lexicon based on WordNet database. It contains more than 65,000 entries. Each entry (synsets) in SWN is given positive, negative and objective scores in the range from 0.0 to 1.0 with overall sum of 1.0. Synset relationship and Gloss description are used to evaluate polarity of entries. Another popular general purpose lexicon is Subjectivity Lexicon (Wilson, Wiebe and Hoffmann, 2005), which is compiled from different resources including General Inquirer (GI), dictionaries and negation files. It consist of more than 8,000 polar (+ve, -ve, and objective) words, with each word being labeled as strong or weak subjective. (Velikovich, Blair-Goldensohn, Hannan and McDonald, 2010) present the development of huge polarity lexicon from web collection. They apply graph propagation algorithm and set of index terms to evaluate polarity of terms with respect to size and quality, with terms taken from web documents. Their lexicon provides sufficient coverage of both positive and negative phrases, covering spelling omissions, and vulgarity in negative sentences expressed in social media posts.

The above mentioned polarity lexicons are used for general purpose opinion mining. However, problem arises when polarity of term changes with the change in domain. It gives rise to development of domain specific polarity lexicons. (Mihalcea, Banea, and Wiebe, 2007) proposed seed words based bootstrapping technique for Romanian language. Their approach is supported by online dictionary and small domain corpus, coupled with filtering module to remove noise from resulting lexicon. Their semi supervised method, experimented on rule based sentence classifier shown better performance over other similar studies. (Bakliwal, Arora and Varma, 2012**)** introduced graph based WordNet propagation method for building Polarity lexicon for Hindi language. They expanded the initial seed lexicon by synonym and antonym relations for adjectives and adverbs only; later on Hindi WordNet is used to assign polarity scores. (Riloff, and Shepherd, 1999,) proposed corpus oriented bootstrapping technique for creating domain specific lexicon. The seed word list is dynamically updated by adding best hypothesis. They applied the algorithm on MUC-4 terrorism domain using eleven semantic classes. Medical opinion mining belongs to specialized domain and requires dedicated medical polarity lexicon. (Goeuriot, Na, Min Kyaing, Khoo, Chang, Theng and Kim, 2012) developed polarity lexicon for health related opinion mining and to the best of knowledge, it is the single study conducted so far on the construction of health reviews based polarity lexicon. Their resulting lexicon is obtained by merging SentiWordNet (SWN) and Subjectivity Lexicon (SL), patient feedbacks are collected to constitute training and testing corpus. All the reviews are labeled as positive and negative with SWN being used for assigning polarity score to each entry. In our work, we tagged each entry of our parsed corpus with UMLS lexicon, then Web thesauri is used to extend the lexicon and finally SWN is used to assign polarity scores for closely matched senses of each medical term. Results of all these studies show that domain specific lexicon helps in improving performance of opinion mining applications.

3. METHODOLOGY

As our approach for constructing medical opinion lexicon is based on mining patient generated reviews, we introduce an incremental modal for creating medical polarity lexicon for medical terms. In each increment, vocabulary of lexicon is enhanced systematically. Its construction is explained below in figure 1.

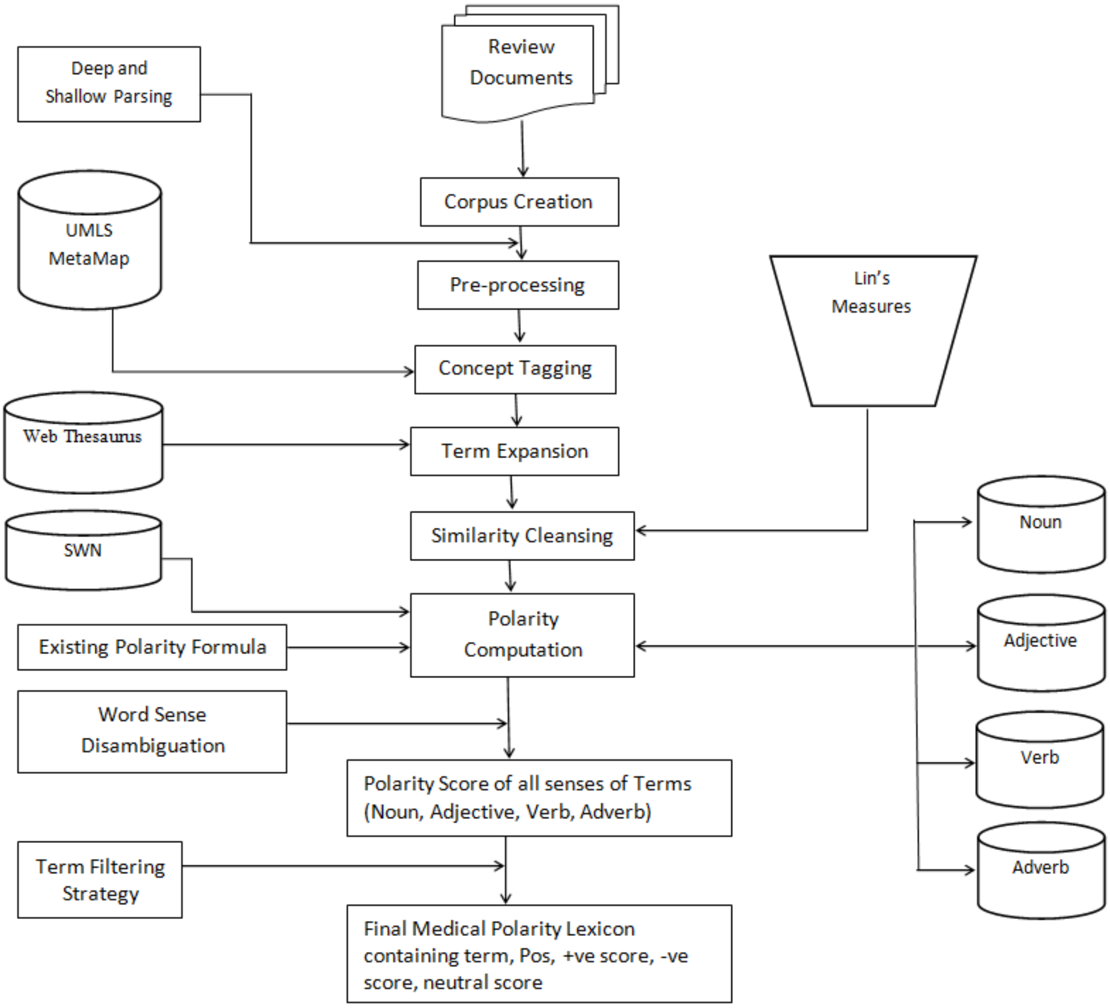

Fig. 1. *Incremental model for medical opinion lexicon generation*

Corpus creation

In this section, we present detailed description of training and testing corpus. As a bench mark corpus for the proposed system is not available therefore we have manually compiled a corpus, representative of such reviews. There are number of websites and forums which contain drug reviews and ratings. Review extraction module of our system dynamically updates review repository because new reviews are regularly posted by users.

Our medical opinion lexicon is constructed using non-experts (patients) drug reviews available on discussion forums, such as Askapatient and DrugRatingz, where users post comments about usefulness and side effects while taking a specific drug. Each review consists of drug name, dosage, comments, patient sex, age and drug rating, example is given below:

Drug: Avelox
Rating: 4
Reason: Sinus Infection
Sex: Female
Age: 38
Comment: It causes unusual headache. It did cure the infection, despite of the side effect.
Dosage & Duration: 400 mg, 10 days (1 x d)

Our corpus contains 30,000 reviews with 2.5 million words for 567 drugs extracted from selected websites. With the help of user specified ratings, 17700(59%) *positive* (rating of 4 or 5), 9900(33%) *negative* (rating of 1 or 2) and 2400(8%) neutral (rating of 3) reviews are identified and stored in a separate text file to constitute the entire corpus. We divide the corpus into training corpus and testing corpus and store them in separate text files.

Training Corpus: It contains 15000 reviews with 7050(47%) positive, 7050(47%) negative and 900(6%) neutral reviews.

Testing Corpus: It contains remaining 15000 reviews, with 8000(53.33%) positive, 5500(36.66%) negative and 1500(10%) neutral reviews.

Pre-Processing

Html parser (Goeuriot, Na, Min Kyaing, Khoo, Chang, Theng and Kim, 2012) is used to extract clean contents into text files by removing html tags and unnecessary metadata information. Sentence boundaries are identified by breaking the cleaned text into sentences. Case-conversion, spelling correction and stop words removal with the help of frequently used stop-word list (Luhn, 1958) is also performed.

In order to get sentence structure, all of the corpus sentences are parsed by using Stanford parser[1]. Parser assign part of speech (POS) tags to each word in sentence. In addition to full parsing, we also perform shallow parsing (Baccianella, Esuli and Sebastiani, 2010) on each sentence to split it into group of related phrases: Noun phrase, verb phrase and prepositional phrase.

Concept Tagging

To check whether a word/phrase in our corpus is a valid medical entry, different medical lexicons are available: Snomed ACT, Mesh, and UMLS for searching drug/medical related content (Erdogan, Halit, Olivier Bodenreider, 2010) Each word in the sentence is searched in UMLS (Bodenreider, 2004), a basic medical lexicon which connects each input word with its corresponding medical concept. UMLS contains more than 1.5 million biomedical concepts and over 10 million associations between these concepts (Nikolova, Ivelina, and Galia Angelova, 2011). We use Sense-Related module of Perl based UMLS similarity package (McInnes, Pedersen, and Pakhomov, 2009) to measure semantic relatedness between input term and its associated concepts listed in UMLS, exactly matched terms are identified and stored along with their UMLS tags and number of occurrences in intermediate lexicon ML-1.

Term Expansion

In next phase, each term of ML-1, previously tagged with health related concept is searched in Web thesauri Thesaurus.com (Kipfer, 1993). (Goeuriot, Na, Min Kyaing, Khoo, Chang, Theng and Kim, 2012) in their work on health related sentiment lexicon construction used Subjectivity Lexicon, but in contrast to their work, we extend our initial lexicon by using Web Thesaurus. We replace subjectivity lexicon with Web Thesaurus (W-T), as this repository contains multiple entries like POS, definition, synonyms and antonyms for a given term. Thus it is more beneficial for lexicon expansion as compared to subjectivity lexicon. This Thesaurus is used to train our medical data to extend the lexicon by including all words within top n entries in treasures.

Filtering

The noise removal from extended lexicon (ML-2) is performed by implementing a filtering step which computes measure of semantic similarity and relatedness between the input term and each of the candidate terms. In our study, we experimented with two measures, namely the Lin's similarity measure (Lin, 1998) and Context vector based semantic relatedness measure (Pedersen, Pakhomov, Patwardhan, and Chute, 2007). In order to stabilize the resulting lexicon, we applied the two measures sequentially; first Lin's similarity score is calculated to filter most nearest terms and then sematic relatedness between two terms is computed as the cosine of angles between two context vectors.

Measuring semantic similarity

The semantic similarity measure depicts likeness of two input terms, and returns a score that determines how much they are identical. It is usually based on is-a relations to link concepts (Lin, 1998; Jiang, Conrath, 1997) available in underlying ontology. For example, smell abnormality and dysosmia are similar in that smell abnormality is a kind of dysosmia in cancer patients. The similarity measures are generally based on path lengths between terms, or they may be based on corpus oriented statistics.

In this work, we have experimented Dekang Lin's proximity based thesaurus (Lin, 1998), which uses a broad set of words in English vocabulary. Each term in the thesaurus has 200 nearest words. The measure of "similarity" is distributional and each of the nearest term receives a similarity score representing the similarity of input term and main term. The Lin's measure provides us a mechanism to filter relevant in extended lexicon by including all the terms within top n entries in Web Thesaurus. It is computed as follow:

$$Lin_{sim} = \frac{2 * IC(LCS)}{IC(c1) + IC(c2)} \quad (1)$$

Equation 1 computes a similarity score between two words senses on the basis of Information Content (IC) of the most specific node (Least Common Subsumer) and the two input concepts (synsets) i.e. c1 and c2. For example, the Lin's linear proximity scores for the input word "Fatigue" and its corresponding treasures are listed in table 1.

In next phase input term of ML-1 along with corresponding treasure from W-T iteratively and calculates Lin's proximity score using Semantic Measures Library[2], which is open source java library for semantic scoring and analysis. This process is repeated until all the input terms of ML-1 are extracted along with their relevant treasures in W-T. The top "n" limit provides us a mechanism to extract relevant words since the lower order words are usually noisy. The terms whose linear proximity value are greater than or equal to 0.2 are included in new lexicon.

[1] nlp.stanford.edu/software/lex-parser.shtml
[2] http://www.semantic-measures-library.org/sml/index.php?q=lib

Table 1. Measuring similarity score for a sample input word

Input Word	Treasure(from W-T)	Lin's Similarity score
Fatigue	Tiredness	1
	Weariness	1
	Lassitude	0.965
	Exhaustion	0.782
	Faintness	0.751
	Fainting	0.453
	Drowsiness	0.432
	Collapse	0.4
	Swoon	0.367
	Shortness of breath	0.207

Measuring Semantic Relatedness

Semantic relatedness is the statistical measure of relatedness between a pair of terms. In literature (Patwardhan, Pedersen, 2006; Patwardhan, 2003; Pedersen, Pakhomov, Patwardhan, and Chute, 2007), different semantic relatedness measures are available such as path finding, information content, and context vectors.

(Pedersen, Pakhomov, Patwardhan, and Chute, 2007) used context vector technique to compute semantic relatedness in biomedical domain. This is intended to be more effective than other measures of semantic relatedness as it is not dependent on path relations among concepts. Our work is an adaptation of this context vector method for dimensionality reduction of the medical opinion lexicon by using corpus of text and intermediate lexicon (ML-2). In this technique, we construct co-occurrence vectors such that each element of the vector contains log-likelihood score between every term present in our corpus and its associated words available in ML-2. The cosine similarity between vectors is then used to determine relatedness of two concepts.

For every term t in our corpus, we create m x n first order co-occurrence vector which stores the occurrence count of every term t. The vector for term t is constructed in three steps, (i) Initialization of all cells of a vector t to zero, (ii) Count the occurrence of each term t in corpus, (iii) For each occurrence of term t; if a term is present in a context window around t then increment by 1 that cell of vector t. The vector is stored in a text file, where each line of the file corresponds to a vector of a term t. In this work, we use manually complied corpus from drug review sites to create term vectors for all content terms occurring in the review collection.

After the creation of term vectors, we then use these to construct context vector corresponding to every UMLS concept whose occurrence in the corpus exceeds a certain limit. The Concept Unique Identifiers (CUIs) are used to identify concepts in UMLS. In addition to CUIs, we also used different relations defined in the UMLS to expand the concepts automatically, such as, parent-child (broader/narrower) relationships. For example "finding" is a parent (broader) and "sign/symptom" is a child (narrower) relationship. (Banerjee and Pedersen, 2003) extended the Lesk's work (Lesk, 1986) by including all the glosses (synsets) for a given term. We use this extended gloss technique as linkedlifedata[3] definition to create a list of index terms for each UMLS concept. The extended gloss terms of linkedlifedata are used to construct context vector for that concept for which there is an exact match between the descriptor term of Web treasure and corresponding concept of UMLS. For example, the UMLS concept "dysgeusia" (UMLS CUI: *C0013378)* maps to cluster of terms in linkedlifedata database.

This cluster also contains gloss terms such as "dysgeusias", "parageusia", "dysgensia", "dysgeusia", "distorted taste", "taste alteration", "taste abnormality", "taste impairment". Therefore, the context vector for "dysgeusia" is represented as: {dysgeusias, parageusia, dysgensia, dysgeusia, distorted taste, taste alteration, taste abnormality, taste impairment}. The formation of these expanded definitions provides a novel contribution of our technique to the existing approaches.

The semantic relatedness of two concepts is calculated as the cosine of the angle between two context vectors:

$$Cos(c1, c2) = \frac{\overrightarrow{v1}.\overrightarrow{v2}}{|\overrightarrow{v1}|.|\overrightarrow{v2}|} \qquad (2)$$

In equation 2 $\overrightarrow{v1}$ and $\overrightarrow{v2}$ are the context vectors for two concepts c1 and c2. The representation of each concept is made by the summation of first order co-occurrence vector for every term. The semantic relatedness of two concepts is the cosine of context vector $\overrightarrow{v1}$ and $\overrightarrow{v2}$. The angle between the two term vectors cannot be greater than $90^{o.}$ The cosine value of two concepts ranges between 0(unrelated) and 1(fully related) and the value closer to 1 means that the medical term is more similar to the compared concept. Selected terms are included in final lexicon (ML-*f*) and irrelevant entries are removed from ML-*f.* The semantic relatedness measurement module is trained on manually constructed corpus consisting of 2.5 million words

3. RESULTS AND DISCUSSION

For each entry in ML-2, our approach automatically retrieves its associated words and their semantic score from SentiWordNet (SWN) (Baccianella, Esuli and Sebastiani, 2010). SentiWordNet is chosen due to its number of distinguishing characteristics including: large word collection (more than 60,000), automatic updates and

[3] http://linkedlifedata.com/resource/umls/id/C0013378

retrieval of each word from WordNet (Wilson, Wiebe and Hoffmann, 2009), and assignment of +ve, -ve and neutral score. As each term in SWN has multiple senses, therefore to compute aggregate score of all senses, we proceed in the following way: for each entry in SWN, positive, negative and objective scores are obtained by computing average of its synsets according to part of speech (POS), i.e., noun, adjective, verb, and adverb using formulas adopted from existing wok (Dang, Zhang and Chen, 2010), i.e.,

Pos_score $(w)_p = \sum_{i=1}^{n} pos\_scorep(i) \;/\; np$
Neg_score $(w)_p = \sum_{i=1}^{n} neg\_scorep(i) \;/\; np$
Obj_score $(w)_p = \sum_{i=1}^{n} obj\_scorep(i) \;/\; np$

Where Pos_score, Neg_score, and Obj_score represent polarity score (positive, negative, objective) of synset i for word w, p denotes part of speech (noun, adjective, verb, and adverb), and n_p is the total number of synsets of the word. For example the word “cold” have thirteen senses under Adjective category and three senses in Noun category We computed the scores a follow:

Pos_score $(\text{“cold”})_{noun}$=0+0+0=0/3=0
Neg_score $(\text{“cold”})_{noun}$=0.125+0.125+0=0.25/3=0.083
Obj_score $(\text{“cold”})_{noun}$=0.875+0.875+1=2.75/3=0.916

The score triplet for word “cold” in noun category is: {0, 0.083, 0.916}. This process is repeated for all POS categories, i.e. adjective, verb, and adverb (in this case adjective only). The word is treated as objective and excluded from final polarity lexicon provided that its objective score is higher than given threshold (0.5). For a word having objective score less than or equal to 0.5, we then checked its positive and negative score. The word with positive score higher than negative score is treated as positive, otherwise it is negative.

The resulting lexicon also contains all of the nearest senses along with polarity score for each medical term as shown in table 2.

Table 2. Final lexicon coverage

Term	Occurrences	Senses
blood pressure	6,713	4
Cold	2,448	16
depression	7,577	3
glucose	11,205	3
ganglion	580	3
discharge	5,072	3
ultrasound	5,704	3
fat	6,112	3
pressure	9,118	4

Evaluation

We manually constructed 15000 reviews with 53.33% positive and 36.66 negative reviews. The algorithms produced polarity lexicon (ML-*f*) which provides sufficient coverage of required data. It contains 2,314 terms, 65% positive and 30% negative and 15% neutral. The most of the terms (73%) stored in ML-*f* are not present in SWN. For example: ultrasound (+ve), blood pressure (-ve), high blood pressure (-ve), Blood pressure determination (neutral), side effect (-ve). Comparison of our lexicons with other two lexicons along with comparing results as shown in table 3 and table 4, which shows the effectiveness of proposed approach

Table 3. Evaluation Results on Training Corpus
(P=precision, R=recall, F=F-score)

	Positive			Negative			Neutral		
	P	R	F	P	R	F	P	R	F
Goeuriot, Lorraine	0.81	0.59	0.68	0.22	0.41	0.29	0.18	0.3	0.2
SWN	0.77	0.41	0.53	0.19	0.41	0.26	0.02	0.15	0.03
M.O.L(this study)	0.82	0.58	0.67	0.24	0.46	0.31	0.16	0.3	0.18

Table 4. Evaluation Results on Testing Corpus
(P=precision, R=recall, F=F-score)

	Positive			Negative			Neutral		
	P	R	F	P	R	F	P	R	F
Goeuriot, Lorraine, et al (2012)	0.76	0.52	0.62	0.24	0.48	0.32	**0.09**	0.04	0.06
SWN	0.74	0.45	0.56	0.22	0.48	0.30	0.05	0.09	0.06
MOL(this study)	0.78	0.56	0.66	0.27	0.51	0.35	0.06	0.09	0.07

4. CONCLUSION AND FUTURE WORK

In this paper, we demonstrated our work on opinion lexicon for health reviews. We have developed a incremental model for creating opinion lexicon with emphasis on health reviews using manually annotated corpus of over 30,000 drug reviews. The evaluation results of experiments reveal that proposed approach gives better results as compared to the other existing methods (e.g. SWN and Goeuriot) and achieves an accuracy of 82% and 74% on training and testing corpuses respectively. Although our technique has shown improved results, automatic expansion of seed cache by using general purpose lexicon such as SentiWordNet and UMLS degrades the results. In future work, further experiments are required to link newly developed lexicon with features of MESH, a rich biomedical resource. Furthermore, neutral reviews also need to be explored and incorporated.

REFERENCES

1. S. Fox. Health information online. The Pew Internet & American Life Project. http://www.pewinternet.org accessed, August 2013.
2. X. Ma, G. Chen, and J. Xiao. Analysis on an online health social network. In Proceedings of the 1st ACM International Health Informatics Symposium, pp 297–306 (2010).
3. A. Jaloba. The club no one wants to join: Online behaviour on a breast cancer discussion forum. *First Monday [Online]*, 14(7): (2009).
4. M.C. Schraefel, R.W. White, P. Andre, D. Tan. Investigating web search strategies and forum use to support diet and weight loss. In Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, pp 3829–3834 (2009).
5. S. Baccianella, A. Esuli, and F. Sebastiani. "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining."Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta, May. European Language Resources Association (ELRA), (2010).
6. P.J. Stone, D. C. Dunphy, M.S. Smith and D.M Ogilvie. The General Inquirer: A Computer Approach to Content Analysis. *Journal of Regional Science*, *8*(1): (1968).
7. T. Wilson, J. Wiebe, and P. Hoffmann. "Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis." Computational linguistics 35(3): 399-433 (2009).
8. L. Goeuriot, L. Goeuriot, J.C. Na, W.Y. Min Kyaing, C. Khoo, Y.K. Chang, Y.L. Theng, and J.J. Kim. Sentiment lexicons for health-related opinion mining. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium. ACM, (2012).
9. L. Velikovich, S. Blair-Goldensohn, K. Hannan, R. McDonald. The viability of web-derived polarity lexicons. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, (2010).
10. T. Wilson, J. Wiebe, and P. Hoffmann. "Recognizing contextual polarity in phrase-level sentiment analysis." Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing. Association for Computational Linguistics: (2005).
11. R. Mihalcea, C. Banea, and J. Wiebe. Learning multilingual subjective language via cross-lingual projections. In Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, CZ, pp 976–983 (2007).
12. A. Bakliwal, P. Arora, and V. Varma. "Hindi subjective lexicon: A lexical resource for hindi polarity classification." (2012).
13. E. Riloff and J. Shepherd. A corpus-based bootstrapping algorithm for semi-automated semantic lexicon construction. Natural Language Engineering, 5(02): 147-156 (1999).
14. H.P. Luhn. The automatic creation of literature abstracts. *IBM Journal of Research and Development*: pp 159-165 (1958).
15. O. Bodenreider. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic acids research*, *32*(suppl 1), D267-D270: (2004).
16. B.T. McInnes, T. Pedersen, and S.VS. Pakhomov. "UMLS-interface and UMLS-similarity: open source software for measuring paths and semantic similarity." AMIA Annual Symposium Proceedings. Vol. 2009. American Medical Informatics Association, (2009).
17. Roget's 21st Century Thesaurus, Third Edition. Retrieved April 06, 2013, from Thesaurus.com website: http://thesaurus.com
18. D. Lin. "An information-theoretic definition of similarity.". In: Proceedings of the 15th International Conference on Machine Learning: Vol. 98. (1998).
19. T. Pedersen, S.V. Pakhomov, S. Patwardhan, and C.G. Chute. Measures of semantic similarity and relatedness in the biomedical domain.*Journal of biomedical informatics*, *40*(3):pp. 288-299 (2007).
20. J.J Jiang, and D.W. Conrath. "Semantic similarity based on corpus statistics and lexical taxonomy." arXiv preprint cmp-lg/9709008 (1997).
21. S. Patwardhan, and T. Pedersen. "Using WordNet-based context vectors to estimate the semantic relatedness of concepts." Proceedings of the EACL 2006 Workshop Making Sense of Sense-Bringing Computational Linguistics and Psycholinguistics Together. Vol. 1501. (2006).
22. S. Banerjee and T. Pedersen. "Extended gloss overlaps as a measure of semantic relatedness." IJCAI. Vol. 3. (2003).

23. M. Lesk. "Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone." Proceedings of the 5th annual international conference on Systems documentation. ACM, (1986).

Y. Dang, Y. Zhang, and H. Chen. "A lexicon-enhanced method for sentiment classification: An experiment on online product reviews." Intelligent Systems, IEEE 25(4): pp. 46-53 (2010).

Erdogan, Halit, Olivier Bodenreider, and Esra Erdem. "Finding Semantic Inconsistencies in UMLS using Answer Set Programming." *AAAI*. 2010.

Nikolova, Ivelina, and Galia Angelova. "Identifying relations between medical concepts by parsing UMLS® definitions." *Conceptual Structures for Discovering Knowledge*. Springer Berlin Heidelberg, Pp. 173-186 (2011).