Procesamiento del Lenguaje Natural, Revista nº 57, septiembre de 2016, págs. 83-90 recibido 28-03-2016 revisado 20-04-2016 aceptado 01-05-2016

COPOS: Corpus Of Patient Opinions in Spanish. Application of Sentiment Analysis Techniques

COPOS: Corpus de Opiniones de Pacientes en Español. Aplicación de Técnicas de Análisis de Sentimientos

Flor Miriam Plaza-del-Arco, M. Teresa Martín-Valdivia, Salud María Jiménez-Zafra, M. Dolores Molina-González, Eugenio Martínez-Cámara
SINAI Research Group
Universidad de Jaén
E-23071 Jaén, Spain
fmpa0002@red.ujaen.es, {maite, sjzafra, mdmolina, emcamara}@ujaen.es

Abstract: Every day more users are interested in the opinion that other patients have about a physician or about health topics in general. According to a study in 2015, 62% of Spanish people access the Internet in order to be informed about topics related to health. This paper is focused on Spanish Sentiment Analysis in the medical domain. Although Sentiment Analysis has been studied for different domains, health issues have hardly been examined in Opinion Mining and even less with Spanish comments or opinions. Thus we have generated a corpus by crawling the website Masquemedicos with Spanish opinions about medical entities written by patients. We present this new resource, called COPOS (Corpus Of Patient Opinions in Spanish). To the best of our knowledge, this is the first attempt to deal with Spanish opinions written by patients about medical attention. In order to demonstrate the validity of the corpus presented, we have also carried out different experiments with the main methodologies applied in polarity classification (Semantic Orientation and Machine Learning). The results obtained encourage us to continue analysing and researching Opinion Mining in the medical domain.
Keywords: Corpus, patient opinions, medical domain, Spanish, Sentiment Analysis, polarity classification

Resumen: Cada día son más los usuarios interesados en la opinión que otros pacientes tienen sobre un médico o sobre temas de salud en general. De acuerdo con un estudio de 2015, el 62% de la población española consulta información en Internet acerca de temas relacionados con la salud. Este trabajo está centrado en el Análisis de Sentimientos en español aplicado al dominio médico. Aunque el Análisis de Sentimientos ha sido estudiado en diferentes dominios, el dominio de la salud apenas ha sido investigado, especialmente en opiniones escritas en español. Por ello, hemos generado un corpus en español con opiniones de pacientes sobre médicos a partir de la extracción de las mismas del portal web Masquemedicos. Este corpus ha sido denominado COPOS (Corpus Of Patient Opinions in Spanish - Corpus de Opiniones de Pacientes en Español). Hasta donde sabemos, es la primera vez que se intenta trabajar con opiniones en español sobre atención médica escritas por pacientes. Para demostrar la validez de este recurso, hemos realizado diferentes experimentos con las principales metodologías aplicadas en la tarea de clasificación de polaridad (Orientación Semántica y Aprendizaje Automático). Los resultados obtenidos nos animan a seguir investigando en el Análisis de Sentimientos en este dominio.
Palabras clave: Corpus, opiniones de pacientes, dominio médico, español, Análisis de Sentimientos, clasificación de polaridad

1 Introduction

The growth of medical documents available on the Internet in the last decade requires the development of more efficient systems to access this kind of information. A 2011 survey of the US population estimated that 59% of

ISSN 1135-5948

all adults have looked online for information about health topics such as a specific disease or treatment (Fox, 2011). Several forums such as Biocreative (Wei et al., 2015), ImageCLEFMed (Müller et al., 2012) or CLEF eHealth (Palotti et al., 2015) have attracted the attention of Natural Language Processing (NLP) researchers (Friedman, Rindflesch, and Corn, 2013). A number of different NLP tasks have been studied in the health domain, from question answering (Lee et al., 2006) to multimodal information retrieval (Martín-Valdivia et al., 2008). Moreover, a number of techniques have been applied to improve the different systems, from basic Machine Learning (ML) algorithms (Chapman et al., 2011) to knowledge integration from medical ontologies (Díaz-Galiano, Martín-Valdivia, and Ureña López, 2009).

Nevertheless, in Sentiment Analysis (SA) it is very difficult to find out about research in the medical field. Although lately the development of SA methods and systems has vastly increased (Liu, 2012), the application of these technologies in the health domain is rather scarce. We can find some recent interesting work which mainly focuses on mining biomedical literature or processing medical web content (Denecke and Deng, 2015). However, most studies only deal with documents written in English, perhaps because platforms for expressing emotions, opinions or comments related to health issues are mainly oriented towards Anglophones. For example, PatientsLikeMe[1] is an online web platform that connect patients with one another, improving their outcomes and enabling research. Other example can be found at the website Patient Opinion[2], in which patients post their point of views after using a health service in United Kingdom, Ireland and Australia.

We consider this is clearly a topic of growing interest not only for people speaking English but also for people who speak a different language, such as Spanish. Our main goal is to launch research in the health domain by mining Spanish patient opinions extracted from the medical web Masquemedicos[3]. Actually, a 2012 survey of the Spanish population estimated that 29.9% of adults have looked online for information about health topics[4]. Nowadays, this number has increased exponentially. According to a study in 2015, 62% of the Spanish people consult Internet to be informed about topics related to the health[5].

In this paper we present the first Corpus Of Patient Opinions in Spanish (COPOS). In addition, we assess the validity of COPOS in implementing two basic polarity classification systems: one based on Semantic Orientation (SO) and another based on ML.

The present paper is structured on the following way: Section 2 describes briefly other studies related to the medical domain. In Section 3 the different resources used and the methodology employed to generate the corpus of opinions of patients are explained. Section 4 shows the experiments carried out and the discussion of the results obtained. Finally, conclusions and future work are presented.

2 Background

As we already stated, research in medical SA is very limited, although we can find some preliminary papers. Perhaps one of the first approaches is that presented by Niu et al. (2005). They manually annotate a corpus of medical abstracts extracted from MEDLINE (1,509 sentences). Then they apply ML (SVM) to classify the polarity of the sentences and the final result is about 79% in recall and precision. Sarker, Molla, and Paris (2011) follow a similar approach but they study the polarity classification at document level over another manually annotated corpus of 520 documents with a total of 9,221 sentences. The system also deals with the detection of negation cues. A comparison with several ML algorithms including SVM, Naïve Bayes, Bayes Net and C4.5 Decision Tree, was carried out and the results are near to 75% in accuracy. Chew and Eysenbach (2010) focus on extracting a corpus from Twitter containing references to the pandemic H1N1 and classify tweets into 16 different categories of opinions and sentiments. Bobicev et al. (2012) also build a corpus of tweets containing Personal Health Information (PHI). They manually annotate the corpus in order to apply ML algorithms to

[1] https://www.patientslikeme.com/
[2] https://www.patientopinion.org.uk/
[3] http://masquemedicos.com/
[4] http://www.ontsi.red.es/ontsi/sites/default/files/informe_ciudadanos_esanidad.pdf
[5] http://insights.doctoralia.es/informe-doctoralia-sobre-salud-e-internet-2015/

classify into positive, negative or neutral the sentiments expressed in the tweets. A similar ML methodology was presented in (Sokolova and Bobicev, 2013), but in this case the corpus used was extracted from a medical forum with messages related to In Vitro Fertilization (IVF). The documents were classified into 5 classes: encouragement, gratitude, confusion, facts, and facts+sentiments. A very interesting point in this paper is the generation of a specific lexicon for the health domain, the HealthAffect Lexicon (HAL). Authors show that the results obtained using HAL are better than applying other general lexicons and features. In a later paper, Bobicev, Sokolova, and Oakes (2015) continue studying the effect of applying HAL over the IVF medical forum, but in this case they focus on analysing sequences of sentiments in online discussions instead of considering only individual posts. This represents a more difficult task oriented towards discourse analysis.

Greaves et al. (2013) apply ML techniques to classifying opinions from patients related to their experience in a hospital of the English National Health Service. They collect a total of 6,412 online comments from patients, also rating the opinions using a scale from 0-5 points. The main goal of the authors was to predict automatically from the textual information in the comment whether the patient would recommend a hospital, whether the hospital was clean and whether he/she was treated with dignity. Another interesting study (Deng, Stoehr, and Denecke, 2014) compares and analyses different lexical and linguistic features in medical documents with sentiments and subjective non-medical texts. The aim is to study the applicability of typical SA methods in clinical narratives. The main conclusion of this study is that a simple method of SA is not suitable for analysing sentiment in clinical documents. Finally, we can find a very good literature review of SA for the medical domain in (Denecke and Deng, 2015).

All the described studies only deal with English documents (Personal Health Information in records or tweets, opinions in blogs or forums). In this paper we focus on Spanish SA in the medical domain. To the best of our knowledge, this is the first attempt to deal with opinions written by patients in Spanish about medical attention. Our approach is similar to the work of (Greaves et al., 2013), but oriented towards applying approaches of SA in the medical forum Masquemedicos. In this site people, mainly without technical or medical knowledge, post opinions and give a ranking for medical entities based on their own experience. Our approach not only applies ML in order to evaluate the viability of our corpus, but we also present a Semantic Orientation method for determining the polarity of the opinions.

3 COPOS: Corpus Of Patient Opinions in Spanish [6]

Due to the growing interest in online patient reviews, we have tried to find a forum or website where opinions are extracted from patients in order to analyze them. Within the medical domain, we have focused on opinions of patients about physicians who they have visited. In choosing the source of information from which to extract the corpus, the following factors were taken into account:

- There must be a reasonable number of opinions and these must be written by patients.
- Each opinion must be assessed by the owner of that opinion.
- The web portal should be a reliable portal in the domain of medicine.
- It must be an internationally prestigious site in search of information about medical entities.

In order to find a source that met all these requirements we conducted an exhaustive study, exploring all possible medical forums containing relevant patient opinions. This task was not easy because there are not too many web sites of patient opinions written in Spanish.

After studying some web sites, our final choice was the medical forum Masquemedicos. The generated corpus is a collection of patient opinions about medical entities that come from six countries (Chile, Colombia, Ecuador, Spain, Mexico, Venezuela). This forum only contains a maximum of 100 opinions per speciality. However, most of the specialities have less than 100 opinions. Moreover, we discarded those opinions that have some empty field. Therefore, the corpus is

[6] http://sinai.ujaen.es/copos-2/

composed of 743 reviews about 34 medical specialities taken on December 3, 2015. Each review contains information about the patient, the medical entity and the textual opinion. About the patient, we obtain his user name or Anonymous (in case of the patient does not show his identity) and his evaluation about the medical entity tagged with stars. In relation to the medical entity, the name and the speciality of the doctor, clinical or hospital are extracted with the city where the consultation was performed. Finally, the textual opinion is composed of positive and negative text parts and the date when the opinion was written. An example of a review of this medical forum can be found in Figure 1.

The reviews are rated on a scale from 0 to 5 stars. A value of 0 means that the patient expresses a very negative opinion about the medical entity, while a score of 5 means that the author has a very good opinion. The number of reviews per rating is shown in Table 1.

Rating	**#Reviews**
0	3
1	88
2	18
3	35
4	51
5	548
Total	**743**

Table 1: Distribution of reviews per rating.

Table 2 shows some interesting features of the corpus. It can be noted that the opinions have an average of 3 sentences, 44 words, 4 adjectives, 3 adverbs, 8 verbs and 10 nouns. We can check that the corpus is completely unbalanced with a portion of positive opinions much higher than negative ones. We have now retrieved all the reviews provided in the Masquemedicos forum. Thus, it seems patients are more interested in writing good comments than bad opinions.

4 *Polarity classification with COPOS*

Polarity classification is one of the most widely studied tasks in SA. This task aims to determine the category of opinion that can be assigned to a text. The category can be binary (positive, negative) or otherwise, and it can be made up of different levels of intensity. In our experiments we consider a binary classification of the reviews of the COPOS corpus (Table 3). In this way, opinions are classified as positive if they have 3, 4 or 5 stars, and negative if their rating is of 0, 1 or 2 stars.

	Positive	**Negative**	**Total**
#Reviews	634	109	743
#Sentences	1,603	406	2,009
#Words	24,244	8,121	32,365
#Adjectives	2,408	594	3,002
#Adverbs	1,695	587	2,282

Table 2: Statistics of COPOS corpus.

Classes	**#Reviews**
Negative	109
Positive	634
Total	**743**

Table 3: Binary classification of COPOS corpus.

Although different approaches have been applied by the research community to tackling the polarity classification task, the mainstream basically consists of two major methodologies: On the one hand, the supervised or ML approach is based on using a dataset to train the classifiers (Pang, Lee, and Vaithyanathan, 2002). On the other hand, the approach based on computing the Semantic Orientation (SO) of the words in the documents does not need prior training, but takes into account the orientation (positive or negative) of words (Turney, 2002). This method is also known as the unsupervised approach. Both methodologies have their advantages and drawbacks. For example, the ML approach depends on the availability of annotated collections of data (training data), and in many cases this is difficult to achieve. On the contrary, a huge amount of lexical resources like lists of opinion words, lexicons or dictionaries, often with dependency on the language, are required by the SO approach. In this paper we present experiments at the document level based on these two methodologies over the COPOS corpus.

In order to tackle the supervised experi-

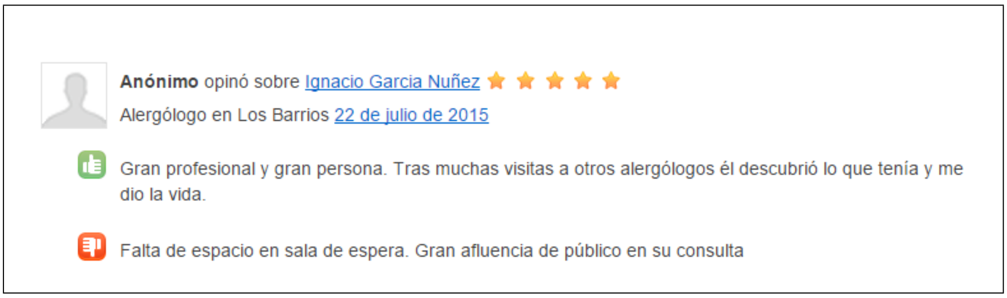

Figure 1: Example of an opinion on the Masquemedicos web portal.

ments the chosen algorithm is the Support Vector Machine (SVM) (Cortes and Vapnik, 1995) because it is one of the most successfully used in Opinion Mining (OM). On the other hand, in the unsupervised experiments we used the improved Spanish Opinion Lexicon (iSOL) (Molina-González et al., 2013), a well-known resource in the SA Spanish community. In order to calculate the polarity (p) of a review (r) with this lexicon, we took into account the total number of positive words (#positive) and the total number of negative words (#negative) within the review, according to the following strategy:

$$p(r) = \begin{cases} 1 & \text{if } \#positive \geq \#negatives \\ -1 & \text{if } \#positive < \#negative \end{cases} \tag{1}$$

where $p(r)$ is the polarity of the review r.

We used the typical evaluation measures employed in text classification: Accuracy (Acc.) and the macro-averaged version of the measures Precision (P), Recall (R) and F1.

4.1 Machine Learning Approach

In this paper we choose the data mining system RapidMiner as a tool for classifying the polarity in the COPOS corpus. The COPOS documents were preprocessed with different combinations of stopper and stemmer in each of the experiments, but in all of them the capital letters were changed to non-capital letters. In order to carry out the supervised approach, the SVM algorithm was applied. The selected SVM algorithm, broadly known by the research community in NLP, uses a linear kernel and normalizes the feature vectors. In order to represent the document we used the TF-IDF weighting scheme. After calculating the features of the documents, a 10-fold cross-validation framework was applied in order to assess the performance of the classifier. The results obtained are shown in Table 4.

Afterward we extended the supervised experiments using a balanced version of the COPOS corpus, formed by 109 negative reviews and 109 positive reviews randomly selected, and the results are shown in Table 5.

4.2 Semantic Orientation approach

The experiments based on Semantic Orientation employed a lexicon-based method. This method involves finding out the presence of opinion words of the lexicon in the documents. As we have mentioned before, the chosen lexicon is iSOL. It is a lexical resource increasingly used for the polarity classification of Spanish reviews. Before carrying out the experiments we performed a preprocessing step on the COPOS corpus in order to apply the same criteria followed during the generation of the iSOL list. For example, for each review we changed capital letters for non-capital letters, accented letters for non-accented letters, and all special characters were deleted from the opinions. Moreover, stop words were discarded.

We first carried out experiments over the original COPOS corpus, classifying a total of 743 reviews. In addition, we also applied our SO approach to the balanced version of COPOS with 218 opinions. Table 6 shows the results achieved by our SO system.

4.3 Result analysis

We consider that the results obtained with the SO approach over the original COPOS corpus are very good, especially taking into account that the results with SVM are similar. In fact, the improvement achieved over the accuracy with the best ML approach is only 0.46%. Bearing in mind that iSOL is a general purpose lexicon, we think that the adaptation of iSOL to the medical domain could achieve very promising results.

With respect to the experiments with the

Stopper	Stemmer	Precision	Recall	F1	Accuracy
YES	YES	78.43%	55.98%	65.33%	86.40%
YES	NO	52.77%	50.83%	51.78%	85.46%
NO	YES	84.17%	**61.40%**	71.00%	**87.88%**
NO	NO	**91.27%**	58.57%	**71.35%**	87.61%

Table 4: Result of polarity classification using SVM over COPOS.

Stopper	Stemmer	Precision	Recall	F1	Accuracy
YES	YES	**90.03%**	**88.32%**	**89.17%**	**88.51%**
YES	NO	88.95%	87.95%	88.46%	88.05%
NO	YES	87.51%	86.09%	86.79%	86.21%
NO	NO	88.23%	86.55%	87.38%	86.69%

Table 5: Result of polarity classification using SVM over balanced COPOS.

COPOS	Precision	Recall	F1	Accuracy
Original	75.28%	70.25%	72.31%	87.48%
Balanced	77.70%	70.66%	74.01%	70.18%

Table 6: Results obtained with iSOL Lexicon over COPOS corpus.

ML method, we first carried out the classification of the reviews with the whole corpus, but due to the highly unbalanced number of opinions for each class (approximately 85% of the opinions are positive and only 15% are negative), we decided to perform a new experiment with a balanced version of the COPOS corpus in order to avoid bias in the classification. However, it is interesting to point out that the results obtained with the balanced corpus are not much better than those with the unbalanced one. It should be noted that with the unbalanced corpus, when a stopper is applied the results are worse while when a stemmer is used the results achieved are better than when it is not applied. On the other hand, when the corpus is balanced the best result is obtained when stemmer and stopper are applied. Experiments conducted on other domains over balanced corpus, such as the movie domain, also obtain the same conclusion (Martínez Cámara et al., 2011).

Regarding the two approaches followed to classify the opinions of COPOS it is noteworthy the difference between the values of F1 and Accuracy measures when the original version of the corpus is used. The main reason for this difference is that Accuracy is a measure that may be biased by the majority class of a dataset. As it was mentioned before, COPOS is a unbalance corpus whose majority class is Positive, which is composed of 634 documents, meanwhile there are 109 negative reviews. Besides, if a classifier reaches a good performance over the majority class of the dataset, there is a higher likelihood that the Accuracy value will be biased. Table 7 shows the confusion matrix of the best configuration when the original version of COPOS is used (Stopper: No; Stemmer: Yes), in which the Precision and the Recall values of the class Positive are very high, meanwhile the performance of the class Negative is not remarkable. Therefore, the difference between the Accuracy and the F1 values are due to the unbalanced nature of COPOS.

5 Conclusions and future work

To the best of our knowledge, we have presented the first Corpus Of Patient Opinions in Spanish (COPOS). In order to demonstrate the usefulness of the corpus, we have carried out experiments with the main methodologies employed in the task of polarity classification (Semantic Orientation and Machine Learning). The results achieved and the growing interest of users in knowing opin-

	True Positive	True Negative	Class Precission
Prediction 1	627	83	88.31%
Prediction -1	7	26	78.79%
Recall	98.90%	23.85%	

Table 7: Confusion matrix of the ML experiment SVM (Stopper: No; Stemmer: Yes) with the original version of COPOS.

ions in the medical domain encourages us to follow this line of research.

Regarding our future work, we plan to extract more patient opinions due to the fact that the corpus presented here is unbalanced. This is an arduous task because in Spanish there are few reliable medical forums with patient opinions. On the other hand, we consider that the adaptation of the general purpose lexicon iSOL to the medical domain would be very interesting research and could greatly improve the final result. Finally, the integration of external medical knowledge, for example extracted from SNOMED, should be investigated.

Acknowledgements

This work has been partially supported by a grant from the Fondo Europeo de Desarrollo Regional (FEDER), REDES project (TIN2015-65136-C2-1-R) from the Spanish Government and by a Grant from the Ministerio de Educación Cultura y Deporte (MECD - scholarship FPU014/00983).

References

Bobicev, V., M. Sokolova, Y. Jafer, and D. Schramm. 2012. Learning sentiments from tweets with personal health information. In *Proceedings of the 25th Canadian Conference on Advances in Artificial Intelligence*, Canadian AI'12, pages 37–48, Berlin, Heidelberg. Springer-Verlag.

Bobicev, V., M. Sokolova, and M. Oakes. 2015. What goes around comes around: Learning sentiments in online medical forums. *Cognitive Computation*, 7(5):609–621.

Chapman, W. W., P. M. Nadkarni, L. Hirschman, L. W. D'Avolio, G. K. Savova, and O. Uzuner. 2011. Overcoming barriers to nlp for clinical text: the role of shared tasks and the need for additional creative solutions. *Journal of the American Medical Informatics Association*, 18(5):540–543.

Chew, C. and G. Eysenbach. 2010. Pandemics in the age of twitter: Content analysis of tweets during the 2009 h1n1 outbreak. *PLoS ONE*, 5(11):e14118, 11.

Cortes, C. and V. Vapnik. 1995. Support-vector networks. *Mach. Learn.*, 20(3):273–297, September.

Denecke, K. and Y. Deng. 2015. Sentiment analysis in medical settings: New opportunities and challenges. *Artificial Intelligence in Medicine*, 64(1):17 – 27.

Deng, Y., M. Stoehr, and K. Denecke. 2014. Retrieving attitudes: Sentiment analysis from clinical narratives. pages 12–15.

Díaz-Galiano, M. C., M. Martín-Valdivia, and L. A. Ureña López. 2009. Query expansion with a medical ontology to improve a multimodal information retrieval system. *Comput. Biol. Med.*, 39(4):396–403, April.

Fox, S. 2011. The social life of health information, 2011. Technical report, PewResearchCenter, May.

Friedman, C., T. C. Rindflesch, and M. Corn. 2013. Natural language processing: State of the art and prospects for significant progress, a workshop sponsored by the national library of medicine. *Journal of Biomedical Informatics*, 46(5):765 – 773.

Greaves, F., D. Ramirez-Cano, C. Millett, A. Darzi, and L. Donaldson. 2013. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *JOURNAL OF MEDICAL INTERNET RESEARCH*, 15.

Lee, M., J. Cimino, H. R. Zhu, C. Sable, V. Shanker, J. Ely, and H. Yu. 2006. Beyond information retrieval — medical question answering. In *Proceedings of* *the AMIA Annual Symposium*, pages 469–473.

Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1):1–167.

Martín-Valdivia, M. T., M. C. Díaz-Galiano, A. Montejo-Raez, and L. A. Ureña López. 2008. Using information gain to improve multi-modal information retrieval systems. *Inf. Process. Manage.*, 44(3):1146–1158, May.

Martínez Cámara, E., M. T. Martín Valdivia, J. M. Perea Ortega, and L. A. Ureña López. 2011. Técnicas de clasificación de opiniones aplicadas a un corpus en español. *Procesamiento del Lenguaje Natural*, 47(0):163–170.

Molina-González, M. D., E. Martínez-Cámara, M. T. Martín-Valdivia, and J. M. Perea-Ortega. 2013. Semantic orientation for polarity classification in spanish reviews. *Expert Syst. Appl.*, 40(18):7250–7257.

Müller, H., A. García Seco de Herrera, J. Kalpathy-Cramer, D. Demner Fushman, S. Antani, and I. Eggel. 2012. Overview of the ImageCLEF 2012 medical image retrieval and classification tasks. In *Working Notes of CLEF 2012 (Cross Language Evaluation Forum)*, September.

Niu, Y., X. Zhu, J. Li, and G. Hirst. 2005. Analysis of polarity information in medical text. In *AMIA 2005, American Medical Informatics Association Annual Symposium, Washington, DC, USA, October 22-26, 2005.*

Palotti, J. R. M., G. Zuccon, L. Goeuriot, L. Kelly, A. Hanbury, G. J. F. Jones, M. Lupu, and P. Pecina. 2015. CLEF ehealth evaluation lab 2015, task 2: Retrieving information about medical symptoms. In *Working Notes of CLEF 2015 - Conference and Labs of the Evaluation forum, Toulouse*, September.

Pang, B., L. Lee, and S. Vaithyanathan. 2002. Thumbs up?: Sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, volume 10 of *EMNLP '02*, pages 79–86. ACL.

Sarker, A., D. Molla, and C. Paris. 2011. Outcome polarity identification of medical papers. In *Proceedings of the Australasian Language Technology Association Workshop 2011*, pages 105–114, Canberra, Australia, December.

Sokolova, M. and V. Bobicev. 2013. What sentiments can be found in medical forums? In *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, pages 633–639, Hissar, Bulgaria, September. INCOMA Ltd. Shoumen, BULGARIA.

Turney, P. D. 2002. Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pages 417–424, Stroudsburg, PA, USA. ACL.

Wei, C.-H., Y. Peng, R. Leaman, A. P. Davis, C. J. Mattingly, J. Li, T. C. Wiegers, and Z. Lu. 2015. Overview of the biocreative v chemical disease relation (cdr) task. In *Proceedings of the fifth BioCreative challenge evaluation workshop.*