Sentiment Analysis of Scientific Citations between Lexical Methods and Statistical Methods

Sara Mifrah*, Oumaima Hourrane*, El Habib Benlahmer*, Nadia Bouhriz*, Mohamed Rachdi*

* *Laboratory of Information Prossessing and Modeling*
Hassan II University of Casablanca faculty of Sciences Ben M'sik
Casablanca, Morocco

DOI: 10.29322/IJSRP.9.12.2019.p9697
http://dx.doi.org/10.29322/IJSRP.9.12.2019.p9697

Abstract- The exponential growth of available scientific articles provides researchers with many new challenges and opportunities. Researchers usually analyses divers' scientific papers to find those pertinent to their research. Citation sentiment analysis is an important task as it can help researchers in identifying shortcomings and detecting problems in a particular approach, determining the quality of a paper for classifying citations that it is positive, negative or neutral in the weighting scheme and recognizing problems that have not been addressed as well as possible gaps in current research approaches. In this paper we present a hybrid approach related to the sentiments analysis generally and citation sentiments analysis specifically; Lexicon-based approaches and Corpus-based approaches (Machine learning and Deep learning). Then, a brief survey study is presented about the use of each approach. Finally a comparison of the used techniques and the ones based on machine learning.

Index Terms- Scientific Citation, Sentiment Analysis, Lexical Methods, Machine Learning, Deep Learning, Word2Vec

I. Introduction

When a scientist produced a writing (article, book, communication to a congress, or other type of document), he must refer to previously published work, because it is an indispensable practice for him. According to [1], each scientific text in a given subject does not exist in itself; it must be included in the literature of this subject.

In a scientific text, the expression used to cite another work is called 'a citation', and the link that indicates the identifier of the cited work is known as 'a reference' [2]. To avoid confusion between these two terms -citation and reference are often used interchangeably-, researchers propose to give definitions to both of them:

Citation : is a call of a reference in a scientific document, it establishes a link with the reference for a defined subject, it is written in a unified way in the text. It is the recognition that a document receives from another [3].

Reference : is the unique identifier of documents at all ranking levels in an act of conference, a book, etc. it is the descrimination that one document gives to another [3].

Citations analysis is an element of bibliometric [4] his goal is : analyze globally using statistical and mathematical methods, the compenents of a documentary corpus to identify the relationships between them. It is an analytical tool that uses citations references from scientific documents, and it is a method of determining the use value of a document, and interpreting the information contained in the citations [5]. It is also a method for obtaining the mechanisms of diffusion and reception of scientific innovations.

To identify the semantic relationship between one document and another, researchers used Sentiment Analysis techniques, which is the task of identifying positive and negative opinions, sentiments, emotions and attitudes expressed in text. Although there has been in the past few years an increasing interest in this field for different text genres such as newspaper text, reviews and narrative text, relatively less emphasishas been placed on extraction of opinions from scientific literature, more specifically, citations.

Citation sentiment detection is an interesting task [6] as it can help researchers in identifying weakness, and detecting problems in a particular approach, determining the quality of a paper, for ranking in citation indexes by including negative citations in the weighting scheme, and recognizing issues that have not been addressed as well as possible gaps in current research approaches. This step comes after another more interesting; it is extracting the citation context.

The extraction of the citations context i.e. where the citations is either in a sentence, or for more precision within a paragraph (it might be interesting to add a few sentences before and after); for example, we can take phrases from the sentence containing the citation to the sentence containing the following citation, or to the end of the paragraph. It is a fundamental step in several applications such as citation summarization, survey generation and citation sentiment analysis.

The purpose of this paper is to present a state of the art of different approaches of sentiment analysis, and its applications in the field of citation analysis

This paper is organized as follows: (Section 2) gives a state of the art of research works that uses sentiment analysis, then the approaches focused on citation sentiment analysis in (Section 3), subsequently (Section 4) present a discussion result of our study, and finally a conclusion (Section 5).

II. Sentiment analysis Approaches

Sentiment analysis or sentiment mining also called polarity mining, opinion mining aims at the analysis of text based on the direction, that is to say a text containing opinions and emotions. The analysis of feelings has generated a great deal of research interest, especially in recent years.

In the literature, sentiment analysis approaches are classified into two categories: Lexicon-based approaches and Corpus-based approaches.

A. Lexical Methods (Lexicon-based Approaches)

These approaches consist of using an external lexicon and dictionaries to extract the sentiment polarity. According to our knowledge, there are some principal work: SentiWordNet with two versions, SentiStrength, The Semantic Orientation CALculator, a holistic approach1 and Serendio approach2. In the follows we will present the first three one.

SentiWordNet is based on WordNet [7][8] which is a lexical database developed by linguists from the cognitive science laboratory at Princeton University for about twenty years. Its purpose is to catalog, classify, and relate in various ways the semantic and lexical content of the English language. WordNet is based on elements called Synsets, which are a set of synonyms and pointers connecting it to other synsets. [9][10] have used SentiWordNet 1.0 and 3.0 respectively which is a lexical resource in which each WordNet synset is associated with three numerical scores Obj (s), Pos (s) and Neg (s), describing how Objective, Positive, and Negative the terms contained in the synset are (Fig. 1).

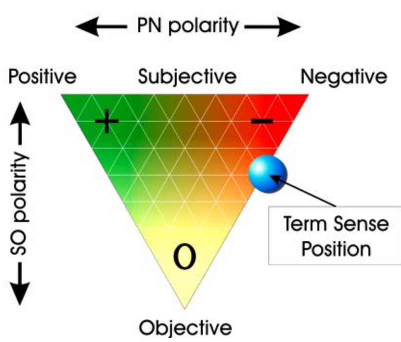

Fig. 1. The graphical representation adopted by SentiWordNet for representing the opinion-related properties of a term sense

A.Esuli and F. Sebastiani [9] based on the quantitative analysis of glosses associated with synsets, and on the use of the resulting vector term representations for semi-supervised synset classification. The three scores are derived by combining the results produced by a panel of eight ternary classifiers, all characterized by similar levels of precision but different classification behaviors. As shown in (Fig.2) the term terrible is provided with two deferent sentiment associations. In this case, SentiWordNet needs to be coupled with a Word Sense Disambiguation (WSD) algorithm to identify the most promising meaning.

Regarding to [10], authors focused on improving the SentiWordNet 3.0 algorithm used to automatically annotate WordNet, they reported the results of the evaluation of SentiWordNet 3.0, against a manually annotated WordNet 3.0 fragment for positivity, negativity and neutrality, and they resulted in accuracy improvements of about 20% over SentiWordNet 1.0.

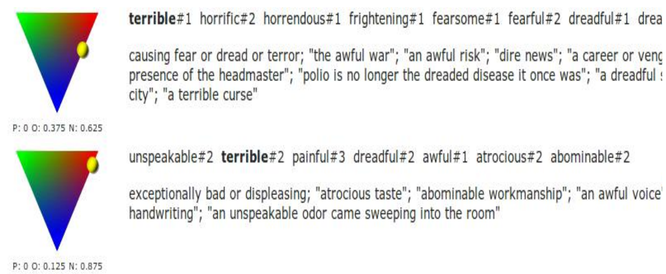

Fig. 1. An example of sentiment association in SentiWordNet

In [11] the authors are useing the available Telugu SentiWordNet to perform sentiment analysis for Telugu e-Newspapers sentences. The proposed system for sentiment analysis has attained an accuracy of 74% for subjectivity classification, and 81% for sentiment classification in the domain of news data.

Other researchers proposes another algorithm SentiStrength [12], which employs several novel methods, to simultaneously extract positive and negative sentiment strength from short informal electronic text. SentiStrength uses a dictionary of sentiment words with associated strength measures, and exploits a range of recognized non-standard spellings, and other common textual methods of expressing sentiment.

Other approach called "The Semantic Orientation CALculator (SO-CAL)" [13] uses dictionaries of words annotated with their semantic orientation (polarity and strength), and incorporates intensification and negation. SO-CAL is applied to the polarity classification task, the process of assigning a positive or negative label to a text that captures the text's opinion towards its main subject matter.

B. Statistical Methods (Corpus-based Approaches)

Unlike the first approach, the second one doesn't use external lexical resources, but it acquires the necessary information to define sentiment from a large corpus. This information is obtained by the application of statistical language models on this corpus.

Sentiment classifiers typically use machine learning [14][15] to identify linguistic features associated with positive and negative sentiments. Such features may include subsets of words, parts of speech [16], n-grams [17][18] (i.e. Successives words), or other lexico-syntactic patterns that signal a polarity positive and negative opinions. These features may be used to train naive Bayes classifiers, which can give probabilistic assignments of positive and negative sentiment to sections of text.

[1] https://www.cs.uic.edu/~liub/FBS/opinion-mining-final-WSDM.pdf

[2] http://www.aclweb.org/anthology/S13-2091

ISSN 2250-3153

In [19] authors are proposed new approach of sentiment classification (Lifelong Learning), it adopts a Bayesian probabilities optimization framework based on stochastic gradient descent. Their experimental results using 20 diverse product review domains demonstrate the effectiveness of the method. They believe that lifelong learning is a promising direction for building better classifier.

Recently, deep learning methods became a popular solution to address sentiment analysis tasks. [20] Studied current works on deep learning for sentiment analysis and grouping deep learning models into recursive, non-recursive, and the mixture of both models. The authors moreover equaled the document level and sentence level sentiment analysis on two datasets. They deduce that deep learning models can be a better solution for polarity detection.

[21] Proposed supervised learning model based on Long Short-Term Memory (LSTM) algorithm. In this model, a sentence representation was made by catching the link between each target word and its contexts. They reported that incorporating target information improves the classification accuracy of the proposed model.

Other researcher they suggested use of deep convolutional neural network (CNN) to extract features from multimodal content and feed these features to a multiple kernel learning classifier for sentiment detection [22]. Their method allow to achieve good results using different datasets.

Regarding to [23] they proposed a hybrid approach that utilizes CNN to learn embedded features and then used a multi-objective genetic algorithm based optimization technique to generate the sentiment augmented optimized vector. They trained Support vector machine (SVM) with non-linear kernel for sentiment detection.

Moreover, the mixture of multiple deep learning algorithms could be useful for improving the performance of sentiment analysis models. [24] Proposed deep learning model by combining CNN with LSTM methods and used word embedding to represent the review sentences. In this model, they used LSTM as pooling layer to capture long term dependencies, which is one of the limitations of CNN algorithm. They evaluated their model using (IMDB) Database and Stanford's sentiment Treebank datasets. The proposed model solved the polarity detection task and mitigate the word order problem. In the same way, [25] combined CNN and LSTM to solve the task of aspect-level sentiment analysis for news articles. This combination is useful when it requires capturing the semantic of the words and the association between them.CITATION SENTIMENT ANALYSIS

In the case of quote, sentiment analysis may be useful in classifying references on a positive–negative scale based on the contexts in which those references occur. General polarities for citations could also be computed by averaging the polarity values of the citation context(s) in which the source is cited. Also, positivity and negativity are sufficient for representing broad patterns of agreement and disagreement among scholars.

Most researchers have used different techniques and tools to extract citation context [26] in a more developed way. For instance, Councill IG et al. [27] have developed ParsCit, which is an open-source software tool for extracting citation context from research papers. This tool utilized conditional random field technique, as a trained machine learning model to annotate the extracted tokens in the references' string, and then identifying and retrieving citation sentences. [28] They have used this tool to parse the entire papers and then extract citation sentences, and other metadata to estimate the importance of each citation.

In [29] authors used SentiWordNet to identify the citation sentiment with the goal of calculating article quality. They extracted all sentences as citation context and they applied part of speech (POS) tags to identify adjectives. They used the SentiWordNet Lexical Analyzer to score each adjective as positive or negative. Lastly, all adjective scores are aggregated to obtain overall sentiment of the sentences.

In [30] the authors are interpreted a sentiment "positive" as an association between the citing author(s) and cited author(s), and a sentiment "negative" as a disassociation between the citing author(s) and the cited author(s). Thus, positive reference contexts may support an author's argument, signal approval of others' work, construct methodology, paytribute, and the like. Negative contexts may function to dispute claims, signal points of disagreement in the domain, correct others' work, and the like. Framing positive and negative polarities in terms of association and disassociation also provides an intuitive mapping from citation contexts to network graphs, which position nodes in relation to each other according to the connections they share (e.g. citing and being cited).

In citation network (Fig. 3), author A who cites author B negatively would still be connected to A, only more distantly than author C who is cited positively by A. If A does not cite D; A would have no connection to author D. The classifier developed in [30], would thus add nuance and sophistication to citation networks, in addition to allowing statistical analysis of various corpora and the citations they contain.

Other approach [31] proposed to detect author's sentiment in the biomedical text documents. This approach based on two steps: (1) classifying the papers as citing or cited papers, and (2) extracting citation sentences from the body of a research citing paper. The researchers applied SVM with kernel function to identify the citation polarity with n-grams as a feature set. SVM with kernel was evaluated on 414 biomedical journals' titles. They classified author's sentiment as positive or others and their model performance was 0.90.

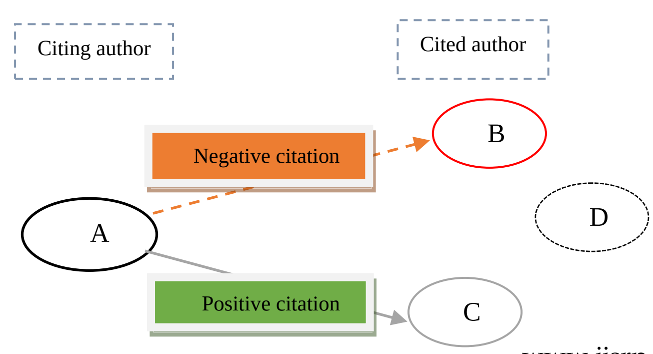

Fig. 2. Example of citation network

In [32], the researchers concentrated on the discussion section of 285 clinical trial papers and developed a rule-based method to extract citation. Furthermore, more than 4000 citations were manually annotated by three annotators. Then, they used SVM with n-grams and sentiment lexicon as a feature set to classify the citation sentences into a positive or negative citation. The combination of their features achieved a better Micro F-score than individual features. Their work developed a good method for citation annotation but the annotation process is conducted manually.

In recent times, Hernandez-Alvarez M and Gomez S JM [33] established annotation methodology for labeling citation sentences to be positive, negative or neutral in order to tackle the citation polarity classification. The authors manually built a corpus and then utilized SVM for citation classification. They tested SVM on labeled corpus and achieved F1 of 0.92.

Finally, Ma et al. [34] they remediated the polarity classification problem by exploiting an author's reputation information. They proposed to use different features, which include author id, affiliation id, polarity distribution, p-index, and unigram. They notified best classification performance through the combination of author id, affiliation id, and p-index feature. Their work was the first survey to improve the traditional citation analysis method for evaluate the quality of scientific research. They shed light on improving H-index method by including negative polarity in the calculation process. However, their work still needs feature engineering skills to use best features for citation sentiment detection.Discussion

The majority of studies for solving citation sentiment analysis issues have used supervised machine learning approaches [35][36][37]. However, these approaches have some difficulties; feature selection process is one of the most important problems that have been used to train the classifiers.

The performance of the classifier depends on the correct selection of the features, on the other hand, there is a difficulty to annotate citation automatically by reason of its special characteristics. Therefore, unsupervised approaches seem to be useful to address some of citation sentiment analysis challenges. For example, distribution representation methods such as word2vec [38][39], and dependency-based word embedding [40], can handle the manual feature extraction in the previous citation sentiment analysis literature. These methods can capture syntactic and semantic structure of citation context.

Although work in this specific area has increased in recent years, there are still open problems that have not been solved, and they need to be investigated. There are not enough open corpus that can be worked in shared form by researchers, there is not a common work frame to facilitate achieving results that are comparable with each other; in order to reach conclusions about the efficiency of different techniques. In this field it is necessary to develop conditions that allow, and motivate collaborative work.

Conclusion

Sentiment can be classified using various categorization schemes and can be expressed at multiple levels of granularity of text. The most popular approach for building a sentiment classifier is using supervised methods in a machine learning framework. However, supervised methods require the availability of labelled data. Labelling this data is a time consuming effort and calls for human annotators, who need to examine and label each data instance. Researchers have thus also explored the use of unsupervised methods for classification of sentiment in different text genres. Most of these unsupervised methods rely on a sentiment lexicon to assign a sentiment score to the text. The sentiment lexicons are largely dependent on the genre of the text being classified. Using the same general purpose lexicon for sentiment analysis in different domains of text is a hard problem. Because there is a marked difference between genres with respect to sentiment detection, most methods are tailored to one particular text type.

Under these variation of techniques that solved the problems of citation analysis, machine learning techniques have been used more than other techniques despite its difficulties.

References

[1] Ziman, J. M. Public knowledge: an essay concerning the social dimension of science. Published by the Syndics of the Cambridge University Press. Bentley House, 200 Euston Road, London, NWI 2DB. American Branch: 32 East 57th Street, New York, N.Y.10022. Library of Congress Catalogue Card Number: 68-10691, ISBN: 0521095190 (1968).

[2] Meyriat, J. Y- a- t- il une place pour une théorie de la documentation ? Revue de Bibliologie- Schéma et Schématisation. n°40., 39-45. (1994).

[3] Hjerppe, Roland. A bibliography of bibliometrics and citations indexing and analysis. Stockholm: the Royal Institute of technology Library. Scientometrics, Vol. 4, Issue 3, 163 p. https://doi.org/10.1007/BF02021064 (1980).

[4] SEMRA, Halima. «Introduction à l 'Analyse de citations : Brève revue de la littérature.». RIST, Volume 11, Numéro 2 (2001).

[5] Garfield, E. "The Science Citation Index as quality information filter," Proceedings of the conference on Coping with Biomedical Literature Explosion: A Qualitative Approach. May 22-23, 1978, NY, New York: The Rockefeller Foundation Working Papers, p.68-77. Proceedings, No:298 (1978).

[6] Awais Athar, Simone Teufel. Detection of implicit citations for sentiment detection. ACL '12 Proceedings of the Workshop on Detecting Structure in Scholarly Discourse, Jeju, Republic of Korea July 12 - 12, 2012. Pages 18-26 (2012).

[7] George A. Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. Introduction to WordNet: An On-line Lexical Database. *International Journal of Lexicography*, Vol. 3, Issue 4, 1 December 1990, 235-244, https://doi.org/10.1093/ijl/3.4.235 (1990).

[8] George A. Miller. WordNet: A Lexical Database for English. Communications of the ACM, November 1995/Vol. 38, No. 11, 39-41, doi: 10.1145/219717.219748 (1995).

[9] Andrea Esuli, Fabrizio Sebastiani. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining. In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06). 417-422 (2006).

[10] Stefano Baccianella , Andrea Esuli , Fabrizio Sebastiani. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10). Valletta, Malta, 2010 may 19-21. Isbn: 2-9517408-6-7 (2010).

[11] Naidu Reddy, Bharti Santosh Kumar, Babu Korra, Sathya Mohapatra and Ramesh Kumar. Sentiment Analysis using Telugu SentiWordNet. International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), SSN College of Engineering, Chennai, India, 22-24 March 2017 (2017).

[12] Thelwall, M., Buckley, K. and Paltoglou, G. Sentiment strength detection for the social web. J. Am. Soc. Inf. Sci., 63: 163–173. doi:10.1002/asi.21662 (2012).

[13] Taboada, Maite, et al. Lexicon-based methods for sentiment analysis. Computational Linguistics 37.2 pp: 267-307. doi: 10.1162/COLI_a_00049 (2011).

[14] Fabrizio Sebastiani. Machine learning in automated text categorization. Journal ACM Computing Surveys (CSUR), Vol. 34, Issue 1, March 2002, pp: 1-47. doi: 10.1145/505282.505283 (2002).

[15] Nasser M. Nasrabadi. Pattern Recognition and Machine Learning," Journal of Electronic Imaging 16(4), 049901 (1 October 2007). http://dx.doi.org/10.1117/1.2819119 (2007).

[16] Paul Schachter, Timothy Shopen. Language Typology and Syntactic Description: Clause Structure. Chapter 1: Parts-of-speech systems. Cambridge University Press. Second Edition - Volume 1, Edited by Timothy Shopen ISBN: 978-0-521-58156-1 (2007).

[17] William B. Cavnar, John M. Trenkle. N-Gram-Based Text Categorization. Available at: http://odur.let.rug.nl/vannoord/TextCat/textcat.pdf (1994).

[18] Peter F. Brown, Peter V. deSouza, Robert L. Mercer, Vincent J. Della Pietra and Jenifer C. Lai. Class-based n-gram models of natural language. Journal Computational Linguistics Volume 18 Issue 4, December 1992, Pages 467-479 (1992).

[19] Chen, Z., Ma, N., & Liu, B. Lifelong learning for sentiment classification. In ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference (Vol. 2, pp. 750-756). Association for Computational Linguistics (ACL) (2015).

[20] Rojas-Barahona LM. Deep learning for sentiment analysis language and linguistics. Compass 10:701–719. https://doi.org/10.1111/lnc3.12228 (2016).

[21] Tang D, Qin B, Feng X, Liu T. Effective LSTMs for target-dependent sentiment classification. In: Paper presented at the The 26th international conference on computational linguistics (COLING 2016). Osaka, Japan., 11–16 Dec 2016, pp 3298–3307 (2016).

[22] Poria S, Chaturvedi I, Cambria E, Hussain A. Convolutional MKL based multimodal emotion recognition and sentiment analysis. In: Paper presented at the 2016 IEEE 16th international conference on data mining (ICDM), Barcelona, Spain, pp 439–448 (2016).

[23] Akhtar MS,Kumar A, EkbalA,Bhattacharyya P. Ahybrid. deep learning architecture for sentiment analysis. In: Paper presented at the the 26th international conference on computational linguistics: technical papers, COLING, Osaka, Japan, 11–17 Dec 2016, pp 482–493 (2016).

[24] Hassan A, Mahmood A. Deep learning approach for sentiment analysis of short texts. In: 2017 3rd international conference on control, automation and robotics (ICCAR), 24–26 April 2017, pp 705–707. https://doi.org/10.1109/ICCAR.2017.7942788 (2017).

[25] Nguyen D, Vo K, Pham D, NguyenM, Quan T. A deep architecture for sentiment analysis of news articles. In: Le N-T, van Do T, Nguyen NT, Thi HAL (eds) Advanced computational methods for knowledge engineering: proceedings of the 5th international conference on computer science, applied mathematics and applications, ICCSAMA 2017. Springer, Cham, pp 129–140. https://doi.org/10.1007/978-3-319-61911-8_12 (2017).

[26] Awais Athar, Simone Teufel. Context-enhanced citation sentiment detection. NAACL HLT '12 Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Montreal, Canada — June 03 - 08, 2012. Pages 597-601. ISBN: 978-1-937284-20-6 (2012).

[27] Councill IG, Giles CL, Kan M-Y. ParsCit: an open-source CRF reference string parsing package. In: Calzolari N (ed) Proceedings of the sixth international conference on language resources and evaluation (LREC'08), Marrakech, Morocco, 28–30 May 2008. European Language Resources Association (ELRA) (2008).

[28] Wan X, Liu F. Are all literature citations equally important? Automatic citation strength estimation and its applications. JASIST 65:1929–1938. https://doi.org/10.1002/asi.23083 (2014).

[29] Sendhilkumar S, Elakkiya E, Mahalakshmi G. Citation semantic based approaches to identify article quality. In: Michal Wozniak HJLD (ed) Proceedings of the third international conference on computer science, engineering and applications-ICCSEA, Delhi, India, 24– 26 May 2013. Springer pp 411–420 (2013).

[30] Chris Alen Sula, Matthew Miller. Citations, contexts, and humanistic discourse: Toward automatic extraction and classification. *Literary and Linguistic Computing*, Volume 29, Issue 3, 1 September 2014, Pages 452 -464. doi: 10.1093/llc/fqu019 (2014).

[31] Kim IC, Thoma GR. Automated classification of author's sentiments in citation using machine learning techniques: a preliminary study. In: Paper presented at the IEEE conference on computational intelligence in bioinformatics and computational biology, CIBCB 2015, 15 Aug 2015, Niagara Falls, ON, Canada, (2015).

[32] Xu J, Zhang Y,Wu Y,Wang J, Dong X, Xu H. Citation sentiment analysis in clinical trial papers. AMIA Annu Symp Proc 2015:1334–1341 (2015).

[33] Hernandez-Alvarez M, Gomez S JM. Citation impact categorization: for scientific literature. In: Paper presented at the IEEE 18th international conference on computational science and engineering (CSE), 21–23 Oct 2015. Los Alamitos, CA, USA, pp 307–313 (2015).

[34] Ma Z, Nam J,WeiheK. Improve sentiment analysis of citations with author modelling. In:KnightK(ed) Proceedings of the fifth workshop on computational linguistics for literature—NAACL-HLT 2016, San Diego, California, USA, 16 June 2016. Association for Computational Linguistics (ACL), pp 122–127 (2016).

[35] Abu-Jbara A, Ezra J, Radev DR. Purpose and polarity of citation: towards NLP-based bibliometrics. In: LucyVanderwendeMR(ed) Proceedings of theNorth American association for computational linguistics (NAACL-HLT 2013), Atlanta, Georga, United States, 9–14 June 2013. Association for Computational Linguistics: Human Language Technologies, pp 596–606 (2013).

[36] Athar A. Sentiment analysis of citations using sentence structure-based features. In: Sasa Petrovic EP (ed) Proceedings of the ACL 2011 student session, Portland, Oregon, 19–24 June 2011. Association for Computational Linguistics, 2000991, pp 81–87 (2011).

[37] Hernandez-Alvarez M, Gomez S JM. Citation impact categorization: for scientific literature. In: Paper presented at the IEEE 18th international conference on computational science and engineering (CSE), 21–23 Oct 2015. Los Alamitos, CA, USA, pp 307–313 (2015).

[38] Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. CoRR Vol. abs/1301.3781, Available at https://arxiv.org/pdf/1301.3781.pdf (2013).

[39] Haixia Liu. Sentiment Analysis of Citations Using Word2vec. CoRR, Vol. abs/1704.00177. https://arxiv.org/abs/1704.00177 (2017).

[40] Wang X, Liu Y, Sun C, Liu M, Wang X. Extended dependency-based word embeddings for aspect extraction. In: Hirose A, Ozawa S, Doya K, Ikeda K, Lee M, Liu D (eds) Neural information processing: 23rd international conference, ICONIP 2016, Kyoto, Japan, 16–21 Oct 2016, Proceedings, Part IV. Springer, Cham, pp 104–111. https://doi.org/10.1007/978-3-319-46681-1_13 (2016).