

# Automated Classification of Author's Sentiments in Citation Using Machine Learning Techniques: A Preliminary Study

In Cheol Kim\* and George R. Thoma

Lister Hill National Center for Biomedical Communications  
National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

**Abstract**— Scientific papers generally include citations to external sources such as journal articles, books, or Web links to refer to works that are related in an important way to the research. The reason for the citation appears within the sentences surrounding the citation tag in the body text, and represents the relationship between the citation and cited works as supportive, contrastive, corrective, etc. This could be an important clue for researchers seeking relevant previous work or approaches for a certain research purpose. We propose to develop an automated method to identify the citing author's sentiments toward the cited external sources expressed in citation sentences using machine-learning techniques and linguistic cues. As a preliminary study, this paper presents a support vector machine (SVM)-based text categorization technique to classify the author's sentiments specifically toward Comment-on (CON) articles. CON, a MEDLINE citation field, indicates previously published articles commented on by authors of a given article expressing possibly complimentary or contradictory opinions. An SVM with a radial basis kernel function (RBF) is implemented, and input feature vectors for the SVM are created based on  $n$ -grams word statistics representing the distribution of words in CON sentences. Experiments conducted on a set of CON sentences collected from 414 different online biomedical journal titles show that the SVM with a RBF yields the best result for an input feature vector combining uni-gram and bi-gram word statistics.

**Keywords**— Citation analysis, author's sentiments, "Comment-on", support vector machine,  $n$ -grams word statistics, MEDLINE

## I. INTRODUCTION

MEDLINE<sup>®</sup> is the premier bibliographic online database of the U.S. National Library of Medicine (NLM) containing more than 24 million citations and abstracts from over 5,600 biomedical journals, and accessed through NLM's PubMed and PubMed Central (PMC) services. With the rapid growth of biomedical literature, both the number of journals indexed and the number of citations produced by NLM are increasing dramatically; 130 journal titles are added each year on average, and nearly 700,000 citations were added to MEDLINE in 2013. The Lister Hill National Center for Biomedical Communications (LHNCBC), a research and development division of NLM, has developed several automated systems that analyze and extract bibliographic information from offline (hard-copy) and online biomedical journal articles to accelerate the production of citation data for MEDLINE, thereby minimizing human labor and providing bibliographic data accurately and in a timely fashion [1].

There are two conventional ways of accessing and navigating the enormous MEDLINE database to get the correct

information: keyword-based searching and tracking citation links between an article and the external sources listed in the reference section through PubMed (or PMC). Researchers may typically use these two methods in combination; first they may try to find representative articles of interest through keyword-based searching and then may collect related works by tracking external sources using citation links provided by PubMed.

However, retrieving relevant articles using these current methods is often time-consuming. PubMed presents users with too many candidates, especially when a search query consists of just a few keywords, or commonly-used or non-specific ones. In addition, PubMed does not provide any further information about the relationship between the articles connected by a citation link. Therefore, researchers need to carefully read the text surrounding each citation tag in the body text of a given article to understand the author's purpose or reason for the citation, thereby purposefully navigating to particular articles or work whose methods and results are in some way related to the given article.

Other highly popular and successful web-based literature searching tools such as Google Scholar and CiteSeer [2] provide not only conventional searching methods based on keywords and citation link information, but also a citation count indicating the number of articles that cite a given article. Thus, users could quite easily search and find works having a high impact on a certain research topic. However, like PubMed, these search tools do not provide the author's reason for citing a particular article or other source.

Generally, authors in the scientific literature include citations to external sources such as journal articles, books, or Web links to refer to works that are foundational in their field, background for their own work, or represent complementary or contradictory research. Authors often mention the reason for a citation within the sentences surrounding the citation tag in the body text. Based on this observation, we propose to develop an automated method for analyzing and identifying the citing author's sentiments toward the cited external sources as expressed within these sentences in the body text. Our method uses machine-learning techniques and linguistic cues.

In this paper, as a preliminary study, we present our automated method using a support vector machine (SVM)-based text categorization technique that classifies the author's sentiments specifically toward Comment-on (CON) articles into two categories: 'Positive' and 'Others', the latter including negative and neutral sentiments. CON is a MEDLINE citation field that indicates a list of previously published articles

commented on by authors of a given article in a complimentary, or sometimes contradictory, manner. We refer to such “Commented on” articles as CON articles, and the papers in which such opinions are expressed as “Comment-in” (CIN) articles. We implemented an SVM with a radial basis kernel function (RBF) as our classifier, and evaluated its performance in terms of accuracy, precision, recall, and F-measure rates. A bag of word-level feature based on  $n$ -grams word statistics representing how differently a word is distributed in ‘Positive’ and ‘Others’ sentiment classes of CON sentences was employed as an input feature vector for the SVM classifier.

## II. CITATION FUNCTION ANALYSIS

“Citation function” has been defined as the citing author’s reason or sentiment toward a cited external source. It therefore represents the relationship between citing and cited works as supportive, contrastive, corrective, etc., and could be an important clue for researchers looking for previous works or approaches for some purpose [3].

Automated analysis of such citation functions is an emerging research topic in the field of natural language processing. It aims to categorize citation functions and to automatically classify citations in scientific literature [4]. Recently, many citation analysis and classification schemes, with a great variance in the number and nature of categories in citation function, have been developed using a variety of text classification methods such as decision tree [5], rule-based method [6], and support vector machines [7]. These citation analysis schemes have now begun to be employed in other areas, such as citation-based text summarization [8], bibliometrics [9], and social media monitoring [10]. Our previous research on Comment-In/Comment-ON (CICO) [11] has also been recognized by other researchers as a first attempt at automatically analyzing citation sentiments for online biomedical articles for MEDLINE, but it is limited to identifying CON citation only in commentary materials [12].

Owing to a wide range of linguistic expressions and writing styles, recognizing an author’s reason for a citation representing the relationship between citing and cited articles is still challenging. Moreover, an author’s reason for including a citation is often not apparent within the text surrounding the citation tag in the body text, and it has been reported that a large proportion of citations is considered just “perfunctory”, i.e., the cited work does not substantially contribute to the citing work [13].

## III. COMMENT-ON SENTENCES

CIN and CON articles are indicated in MEDLINE fields as “Comment in” and “Comment on” respectively, and are linked together. As an example, Fig. 1(a) is the MEDLINE citation for an article (CIN) in which a “Commented on” article is cited. This CON information, shown enclosed in a dotted box, consists of the abbreviated journal title, publication year, volume and issue number, and pagination. Conversely, as shown in the dotted box in Fig. 1(b), the MEDLINE citation for this CON article cites the CIN article in which it is mentioned. Thus readers may get to either citation from the other.

In an article, a sentence associated with a citation tag (such as “(1)” or “[1]”) that points to the complete bibliographical description of the cited external source listed in the reference

section is called a “citation sentence”. In this study, we also define a “CON sentence” as a citation sentence that specifically indicates a CON article. CON sentences are therefore a subset of citation sentences.

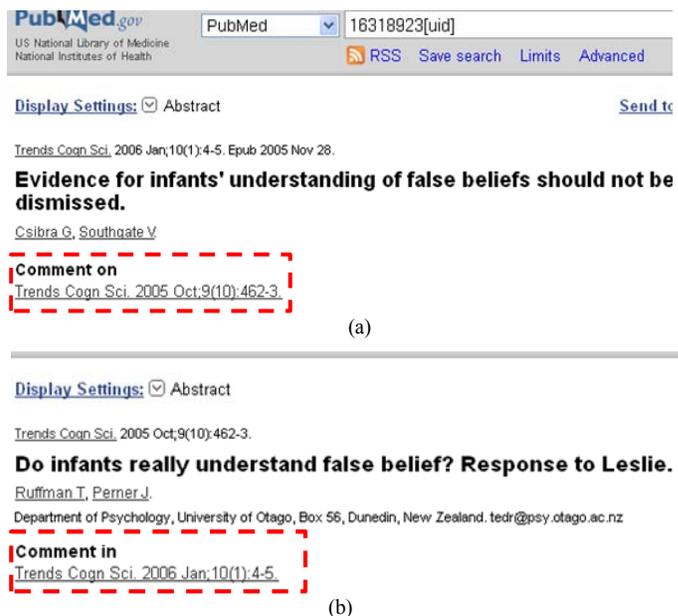


Fig. 1. (a) “Comment on” and (b) “Comment in” citations in MEDLINE

CIN articles are usually short papers such as commentaries, letters, editorials, or brief correspondences, written mainly for the purpose of supporting, refuting, or discussing other articles (CON); authors of a CIN article cite CON articles related to their research as primary external sources. Accordingly, a CON sentence is very likely to include evidence of the author’s sentiment (complimentary or contradictory), and a concise description of the methods or findings reported in the CON article. Based on such observation and analysis, we define three categories of citation functions (author’s sentiments) as positive, negative, and neutral. Here, ‘neutral’ represents the citing author’s objective description of the cited work (neither positive nor negative). Typical examples of CON sentences in each category of citation function are shown in Table 1.

While building the ground-truth dataset for training and testing, we found that negative CON sentences are relatively rare when compared to the CON sentences having other sentiments, thereby heavily skewing the distribution of CON sentences in each sentiment class. Earlier studies [4][7] have suggested that this might be because negative sentiments could be politically dangerous, and thus authors tend to express these in a more subtle manner. For example, authors often express their negative views toward an external source not in the corresponding citation sentence but rather in other sentences located right before or after it.

In order to test our idea simply but reliably, using our CON dataset which has a highly uneven sentiment distribution, we merge CON sentences in the negative and neutral classes into one class, called ‘Others’. As a result, our task of identifying the citation sentiment in a CON sentence is redefined as a two-class problem, thereby classifying a given CON sentence as either

‘Positive’ or ‘Others’.

TABLE I. EXAMPLES OF CON SENTENCES IN EACH CATEGORY OF CITATION FUNCTION .

Sentiments	CON sentences
Positive	We read with interest the very intriguing case reported in Reproductive Toxicology by Kim et al. [1].
	I enjoyed the article by Cooper et al [1] and was delighted to see some evidence being published outlining the role of the emergency care practitioner (ECP).
	Fleming et al [1] are to be commended for the excellent technical presentation of portal vein reconstruction using clear art work and intraoperative photographs.
	We would like to compliment Hoilund-Carlsen and colleagues on their well-designed study on myocardial perfusion scintigraphy (MPS) as gatekeeper for coronary angiography. [1]
Negative	We are gravely concerned that the conclusions reached by Bandak [1] may be invalid due to apparent numerical errors in his estimation of forces experienced in an infant neck during vigorous shaking.
	It is unfortunate that Doctor Hall and colleagues have not referenced these controversies in their otherwise excellent review article. [7]
	In the meta-analysis by Miller et al. [1], the study design, data analysis, and main conclusions seem to have substantial drawbacks and to be affected by poorly controlled clinical and statistical variables.
	Editor—We disagree with Luty et al's suggestion that burpenorphine should replace methadone. [1]
Neutral	Pascual and colleagues studied the impact of pretreatment with statins on patients scheduled to undergo coronary artery bypass surgery (CABG) [1].
	The paper by Shiryayev et al. [8] published in this issue of Biochemical Journal aimed to investigate the catalytic properties of NS3 from WNV and to identify potent inhibitors.
	The recent study by Dick et al [1] suggests that male patients with a severe carotid stenosis are at a higher risk of vascular events (mainly stroke) compared with women.
	It was recently described in Cancer Research by Laurent et al. [1] how low molecular weight superoxide dismutase (SOD) mimetics increase H <sup>2</sup> O <sup>2</sup> levels, which in turn killed colon (CT26) and liver (Hepa 1-6) tumor cells.

#### IV. PROPOSED METHOD

In this paper, we present an automated method for classifying the author’s sentiment expressed in CON sentences using an SVM-based text categorization technique and a bag of words input feature based on  $n$ -grams word statistics. Our method consists of four main steps: 1) classification of an online biomedical paper as either a CIN (letter-like short paper) or a regular full-length article, 2) detection and extraction of citation sentences from the body text of a given CIN article, 3) identification of CON sentences from a set of citation sentences, and 4) classification of author’s sentiment for a citation from each CON sentence by the SVM. We accomplish steps 1) to 3) using machine-learning based methods developed in our previous studies [11][14][15]. Here, we focus on step 4), analyzing and classifying author’s sentiments expressed within CON sentences.

##### A. Feature Extraction

In our research, we adopt a bag of words based on  $n$ -grams—specifically, uni-gram ( $n = 1$ ) and bi-gram ( $n = 2$ )—of word statistics representing how differently a word or a pair of words is distributed in ‘Positive’ and ‘Others’ sentiments of CON

sentence classes, to build an input feature vector for the SVM classifier. Using words as input features requires a very high dimensional feature space (10,149 dimensions of uni-gram in our case). Although the SVM can manage (lead to a convergence) such a high dimensional feature space, many have suggested the need for word selection or dimension reduction to employ other conventional learning methods, reduce the computational cost, improve the generalization performance, and avoid the over-fitting problem. A typical approach for word selection is to sort words according to their importance. Many functions have been proposed to measure the importance of a word, including term frequency (TF), inverse document frequency (IDF),  $\chi^2$  statistics, and simplified  $\chi^2$  ( $s\chi^2$ ) statistics [16]. The use of  $s\chi^2$  has been reported as delivering the best performance since it removes redundancies, and emphasizes extremely rare features (words), and rare categories from  $\chi^2$  [17].

In our task,  $s\chi^2$  of word  $t_k$  for CON sentences in the ‘Positive’ sentiment class (class  $c_0$ ) and those in the ‘Others’ class (class  $c_1$ ) can be defined as follows:

$$s\chi^2(t_k, c_i) = P(t_k, c_i) \cdot P(\bar{t}_k, \bar{c}_i) - P(t_k, \bar{c}_i) \cdot P(\bar{t}_k, c_i) \quad i = 0, 1 \quad (1)$$

where  $P(t_k, c_i)$  denotes the probability that, for a random sentence  $x$ , word  $t_k$  occurs in  $x$ ,  $x$  belongs to class  $c_i$ , and is estimated by counting its occurrences in the training set. The importance of word  $t_k$  is finally measured as follows:

$$s\chi_{max}^2(t_k) = \max_i s\chi^2(t_k, c_i) \quad i = 0, 1 \quad (2)$$

Accordingly, the more differently a word is distributed in ‘Positive’ sentiment and ‘Others’ classes, the higher its  $s\chi_{max}^2(t_k)$ . Words are sorted according to their  $s\chi_{max}^2$  and a word dictionary that is created by selecting words having highest  $s\chi_{max}^2$  scores is then considered a bag of words feature. Table 2 shows lists of the top 20 uni-gram and bi-gram words based on their  $s\chi_{max}^2$ . Finally, the bag of words feature is converted to a binary vector for SVM: The vector dimension corresponds to the number of words in the dictionary, and each vector component is assigned 1 if the corresponding word in the dictionary is found in a given CON sentence or 0 otherwise. We performed a series of experiments to investigate the influence of word reduction and to discover the best-performing word dictionary size.

TABLE II. LISTS OF TOP 20 UNI-GRAM AND BI-GRAM WORDS SCORING HIGHEST  $s\chi_{max}^2$

<b>Uni-gram</b>	Important, interest, read, article, editor, not, interesting, Provide, issue, new, great, patients, et, provides, further, al, novel, dr, excellent, paper
<b>Bi-gram</b>	read with, with interest, with great, great interest, associated with, did not, agree with, have read, very interesting, compared with, our understanding, patients with, interesting article, has been, first time, important study, further evidence, important new, excellent article, new insights

### B. SVM Classifier

SVM [18] was originally introduced as a supervised learning algorithm based on the structural risk minimization principle for solving a two-class problem, though it can be easily extended to handle multi-class problems. Owing to its consistently superior performance compared to other existing methods, SVM has been widely used in many text categorization and summarization tasks. The basic idea of using SVM to solve a non-linear pattern recognition problem is to map a non-linear separable input space to a linear separable higher dimensional feature space using a predefined kernel function, and to find the optimal hyperplane that maximizes the margins between the classes in that feature space.

We employed a SVM with a radial basis kernel function (RBF) defined in equation (3) below which has been commonly used in pattern recognition applications, and implemented it using LibSVM, a free software package for non-commercial use [19].

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (3)$$

## V. CLASSIFICATION EXPERIMENTS

### A. Database

In order to build a ground-truth dataset for our experiments to automatically categorize author’s sentiments in CON sentences, we first collected 2,665 CON sentences from online biomedical articles published in 414 different journal titles and indexed in MEDLINE. As mentioned previously, these online articles are letter-like short papers, and their publication types are Letter (49.0%), Review (2.1%), Editorial (25.4%), Commentary (14.5%), and others (9.0%).

The collected set of CON sentences are then divided into three classes (‘Positive’, ‘Negative’, and ‘Neutral’) according to the author’s expressed opinion of the citation within those CON sentences through a manual annotation process. Among these, 2,109 CON sentences consisting of 936 in the ‘Positive’ sentiment class and 1,173 in the ‘Others’ (‘Negative’ + ‘Neutral’) class were randomly selected to train the SVM. The statistics ( $s\chi_{max}^2$ ) of  $n$ -grams in the CON sentences are also estimated from this training set. The remaining 556 sentences (255 from the ‘Positive’ class + 301 from the ‘Others’ class) were used as a test set to evaluate the performance of the SVM.

### B. Experimental results

In experiments, we evaluated the performance of our SVM classifier in identifying the author’s sentiment for a citation in CON sentences in terms of accuracy, precision, recall, and F-measure rates that are defined as follows;

$$accuracy = (TP + TN)/(TP + FP + TN + FN)$$

$$precision = TP/(TP + FP)$$

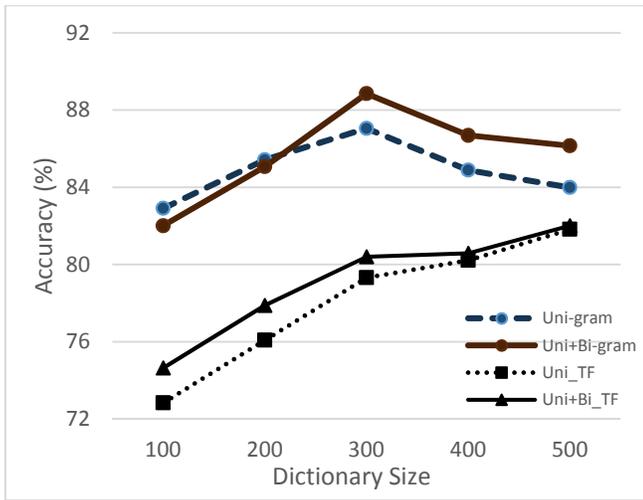
$$recall = TP/(TP + FN)$$

$$F\text{-measure} = 2 \times (precision \times recall)/(precision + recall)$$

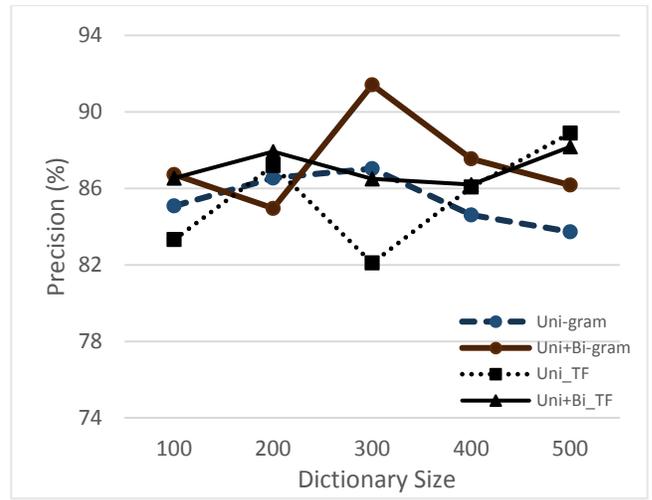
Here,  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  denote true positive, true negative, false positive, and false negative, respectively. False negative means that the ‘Positive’ sentiment in a CON sentence is misclassified into the ‘Others’ class. False positive is the reverse of the above.

Figure 2 shows the accuracy, precision, recall, and F-measure rates of the SVM as functions of the size of the word dictionary in the bag of  $n$ -grams ( $n = 1, 2$ ) words features. As mentioned earlier, words in these bag of words features are selected according to their corresponding  $s\chi_{max}^2$  scores that reflect the difference of their distributions between the ‘Positive’ and ‘Others’ classes. It can be seen that our SVM classifier performs better when the combination of uni-gram and bi-gram, rather than uni-gram alone, is used as an input feature vector, especially when the size of the word dictionary is 300 or more.

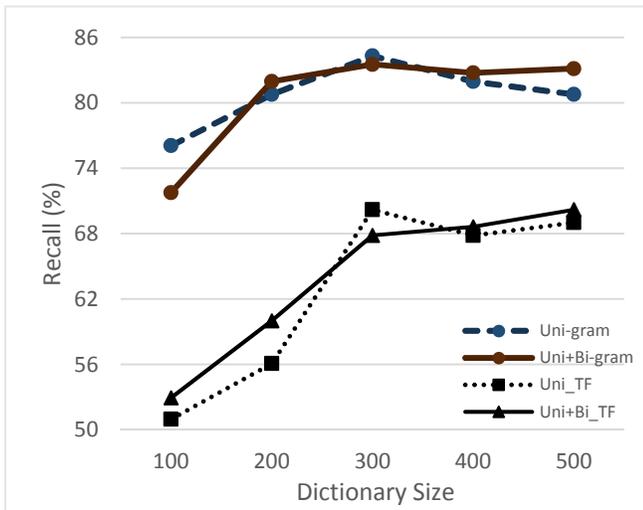
For a comparison study, we also employed and tested another bag of words input features created using the term frequency (TF), which were employed in earlier citation analysis studies [7][20]. In those features, a set of  $n$ -gram words is simply gathered based on the number of their occurrences in the training dataset. As can be seen in Fig. 2, the SVM with a RBF yields a remarkably better performance overall when a bag of words feature is created based on  $s\chi_{max}^2$ , compared to that created using TF.



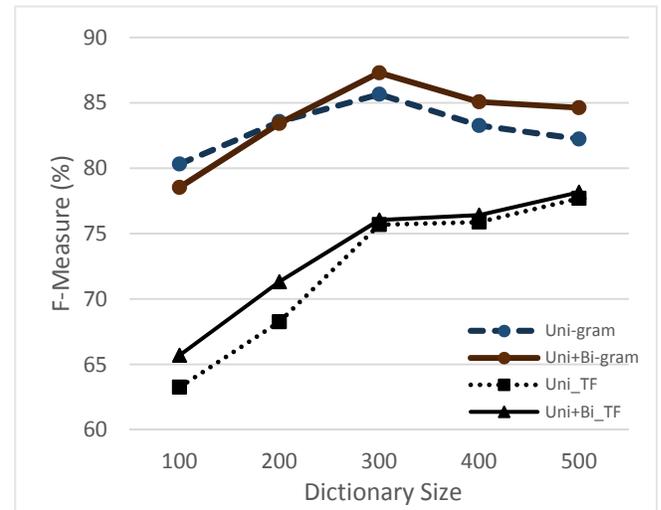
(a)



(b)



(c)



(d)

Fig. 2. (a) Accuracy, (b) Precision, (c) Recall, and (d) F-Measure rates of SVM with a RBF against different word dictionary sizes in the bag of  $n$ -grams (Uni-gram and Bi-gram by  $s\chi^2_{max}$ , and Uni\_TF and Bi\_TF by TF).

TABLE III. THE RESULT OF EXTRACTING "CITATION SENTENCES" USING THE SVM WITH A RBF

Error types	CON sentences
False-Negative	<p>This is exactly what de Jonge et al. [4] achieved with their <b>groundbreaking</b> investigation presented in this issue of the Journal.</p> <p>Thus, the report of Assmus et al., [9] in this issue of the Journal, showing that intracoronary infusion of BMC did <b>not</b> aggravate restenosis development <b>nor</b> was associated with an increase of cardiovascular events, including the necessity for repeated coronary revascularization procedures, is very <b>reassuring</b>.</p>
False-Positive	<p>We read with interest the impressive meta-analysis by Davis et al [1] of the efficacy of second-generation antipsychotics (SGAs) published in the ARCHIVES but were concerned about its inadequate consideration of some important methodologic limitations that may have significantly detracted from the veracity of their conclusions.</p>

Table 3 shows examples of false-negative and false-positive classification errors from the SVM. The first CON sentence in the false-negative error examples contains the word, ‘groundbreaking’ which suggests a positive sentiment. However, this word is missing in the word dictionary of the bag of words input features, certainly due to the small size of our current training dataset. Thus we expect that this type of errors can be fixed by collecting more CON sentences and increasing the size of the training dataset. The second false-negative error in Table 3 is misclassified due to negated words (‘not’ and ‘nor’) even though it also has a positive word (‘reassuring’) expressing the author’s real sentiment.

In citation sentences, negative sentiment is often expressed in subtle ways or mitigated by starting with praise. The example of a false positive error shown in Table 3 is typical. While the first half of the sentence praises some aspects of the cited paper, the remaining part describes the citing author’s concern about its shortcomings. Clearly, criticism is the intended sentiment. Such a subtle citing manner makes the problem of recognizing the author’s negative opinions often very challenging.

## VI. FUTURE WORK

In future research, we first plan to improve the performance of our proposed method for classifying author’s sentiments in CON sentences. In this preliminary study, through a series of experiments and error analysis, our ground-truth training dataset was found not big enough to reliably calculate  $n$ -grams word statistics employed to create the bag of words input features for the SVM classifier. Consequently, the author’s sentiment in a given CON sentence is occasionally misclassified despite the

fact that positive or negative meaning of words clearly exist within the sentence. In order to minimize this problem, we are considering a significant increase in the size of the ground-truth training dataset by collecting more CON sentences, though a time-consuming manual annotation process is also required.

On the other hand, we learned that the authors’ reason for citing often does not appear clearly within the corresponding citation sentence, especially when they criticize some aspects in the cited work. Rather, other sentences located right before or after the citation sentence are found to have better linguistic clues or contextual information about the author’s intended reason for citing. As an example, two citation sentences that are shown in bold text in Table 4 seem to have no explicit expressions of author’s sentiments on the citations. However, from the sentences right after these citation sentences, we can easily recognize author’s intended sentiment (positive for the first citation and negative for the second). Therefore, if we focus solely on the citation sentence, we may lose a significant amount of clues suggesting the author’s real sentiments. To deal with such a problem, we plan to extend the range of the text of interest for searching for relevant context representing the true sentiment of the author towards the cited paper. This text of interest, called “citation sentence+”, can be as short as a single sentence or span across multiple sentences. We will also develop a reliable method to determine the range of such text of interest.

In addition, other types of input features, including surface-level features such as sentence position within the body text, and similarity of titles between the citing and cited articles will be employed and tested to compensate for errors from the current bag of words features based on  $n$ -grams word statistics.

TABLE IV. EXAMPLES OF “CITATION SENTENCE+”

Sentiments	Text
Positive	<b>We would like to comment on the retrospective review of melioidosis by Chan et al in a recent issue of CHEST (November 2005).</b> <sup>1</sup> We commend the authors on their work, which reinforces the high mortality rate associated with this infection, particularly in those patients with a critical illness.
Negative	<b>We would like to comment on the article by Hurwitz et al<sup>1</sup> published recently in the Journal of Clinical Oncology.</b> This article states that fluorouracil plus leucovorin plus bevacizumab seems as effective as irinotecan plus fluorouracil plus leucovorin. We think that this article, presented as a formal phase III, in a prestigious journal can deeply mislead the reader.

## VII. CONCLUSIONS

Authors in the scientific literature generally include citations to external sources in their papers to refer to works that are related in an important way to their research. The author’s reason or sentiment toward a citation, which usually appears within the sentences surrounding the citation tag in the body text, is called the citation function. Thus citation function represents the relationship, such as supportive, contrastive, corrective, etc., between citing and cited works, and could be an important clue for researchers, particularly those who are seeking previous work or approaches for some research purpose.

In this preliminary study, we have presented an automated

method using a support vector machine (SVM)-based text categorization technique that classifies the author’s sentiments toward cited work into two categories: ‘Positive’ and ‘Others’. CON is a MEDLINE citation field showing previously published articles commented on by authors of a given article (“Comment-in” or CIN) as primary external sources on which they may express complimentary or contradictory opinions. We have implemented an SVM with a radial basis kernel function (RBF) as a classifier and evaluated its performance in terms of accuracy, precision, recall, and F-measure rates. A bag of words based on  $n$ -grams word statistics that represents how differently a word is distributed between ‘Positive’ and ‘Others’ sentiment classes of CON sentences is employed to build an input feature

vector for the SVM.

Through a series of experiments on a set of CON sentences collected from 414 different online biomedical journal titles, we see that the SVM with a RBF yields the best performance overall (around 90%) when the bag of words input feature is based on a combination of uni-gram and bi-gram, and its word dictionary size is 300. A comparison study also shows that the bag of words input features in which words are selected based on  $n$ -grams word statistics performs remarkably better than those based on simple term frequency.

Error analysis also suggests future research to improve the overall performance of classifying the author's sentiment in a CON sentence by: 1) increasing the size of the ground-truth dataset, 2) employing more input features and testing other types of machine-learning techniques, and 3) developing a reliable method for determining the range of a text of interest, called "citation sentence+" that surrounds a citation.

#### ACKNOWLEDGMENT

This research was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

#### REFERENCES

- [1] G.R. Thoma, D.X. Le, I. Kim, J. Kim, C. Moon, L. Tran, and J. Zou, "Automation to accelerate the production of MEDLINE," *Technical Report to the LHCBC Board of Scientific Counselors*, April 2008.
- [2] C. L. Giles, K. D. Bollacker, and S. Lawrence, "CiteSeer: an automatic citation indexing system," *Proc. the Third ACM Conf. Digital Libraries*, pp. 89-98. ACM Press, 1998.
- [3] S.B. Shu, "Evolving the web for scientific knowledge: first steps towards an HCI knowledge web," *Interfaces, British HCI Group Magazine*, vol. 39, pp. 16-21, 1998.
- [4] S. Teufel, A. Siddharthan, and D. Tidhar, "Automatic classification of citation function," *Proc. the 2006 Conf. Empirical Methods in Natural Language Processing (EMNLP 2006)*, pp.103-110, 2006.
- [5] R.E. Mercer and C. Di Marco, "A design methodology for a biomedical literature indexing tool using the rhetoric of science," *Proc. Workshop on Linking Biological Literature Ontologies and Databases*, pp. 77-84, 2004.
- [6] U. Schafer, and C. Spurk, "TAKE scientist's workbench: semantic search and citation-based visual navigation in scholar papers," *Proc. the 2010 IEEE Fourth Int'l Conf. Semantic Computing (ICSC)*, pp. 317-324, 2010.
- [7] A. Athar, "Sentiment analysis of citations using sentence-structure-based features," *Proc. NAACL-HLT 2012*, pp. 597-601, Portland, June 2011.
- [8] A. Abu-Jbara and D. Radev "Coherent citation-based summarization of scientific papers," *Proc. the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 500-509, 2011.
- [9] A. Abu-Jbara, J. Ezra, and D. Radev, "Purpose and polarity of citation: towards NLP-based bibliometrics," *Proc. NAACL-HLT 2013*, pp. 596-606, Atlanta, June 2013.
- [10] N. Godbole, M. Srinivasaiah, and S. Skiena, "Large-scale sentiment analysis for news and blogs," *Proc. the 9th Int'l Conf. Weblogs and Social Media (ICWSM-15)*, Boulder, March 2007.
- [11] I. Kim, D.X. Le, G.R. Thoma, "Identifying Comment-on Citation Data In Online Biomedical Articles Using SVM-based Text Summarization Technique," *Proc. Int'l Conf. Artificial Intelligence*, pp. 431-437, Las Vegas, 2012.
- [12] B. Yu, "Automated citation sentiment analysis: What can we learn from biomedical researchers," *Proc. The American Society for Information Science and Technology (ASIST)*, Vol. 50, no. 1, pp. 1-9, 2013.
- [13] M.J. Moravcsik and P. Murugesan, "Some results on the function and quality of citations," *Social Studies of Science*, vol. 5, pp.88-91, 1975.
- [14] I. Kim, D.X. Le, and G.R. Thoma, "Automated identification of biomedical article type using support vector machines," *Proc. 18th SPIE Document Recognition and Retrieval*, 7874:787403, pp. 1-9, San Francisco, January 2011.
- [15] I. Kim, D.X. Le, and G.R. Thoma, "Automated method for extracting "citation sentences" from online biomedical articles using SVM-based text summarization technique," *Proc. the 2014 IEEE Int'l Conf. on Systems, Man, and Cybernetics (SMC 2014)*, pp. 2006-2011, San Diego, October, 2014.
- [16] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.
- [17] L. Galavotti, F. Sebastiani, and M. Simi, "Experiments on the use of feature selection and negative evidence in to automated Text categorization," *ECDL 2000 LNCS 1923*, pp. 59-68, Springer, 2000.
- [18] ] V. Vapnik, *The nature of statistical learning theory*, New York: Springer-Verlag, 1995.
- [19] C.C. Chang and C.J. Lin, "LIBSVM: a library for support vector machines," 2001. [<http://www.csie.ntu.edu.tw/~cjlin/libsvm>]
- [20] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," *Proc. ACL-02 Conf. Empirical methods in natural language processing (EMNLP '02)*, vol. 10, pp.79-86, Stroudsburg, 2002.