Machine Learning with Selective Word Statistics for Automated Classification of Citation Subjectivity in Online Biomedical Articles

Incheol Kim* and George R. Thoma
Lister Hill National Center for Biomedical Communications
National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

Abstract - *There is growing interest in automatically classifying author's sentiment expressed within citation sentences in scientific literature to provide effective tools for researchers who are seeking relevant previous work or approaches for a certain research purpose. We propose an automated method of determining whether a given citation sentence contains an author's subjective opinion (positive or negative) or objective factual information, as the first step to analyze and identify the citing author's sentiments toward the cited external sources. Our method uses a support vector machine (SVM)-based text categorization technique to identify the subjective citations 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. We introduce a bag of unigrams based on selective word statistics, which is derived from a text region of interest within a sentence containing a description of author's reason of citation and lexical linguistic cues to build an input feature vector for the SVM classifier. Experiments conducted on a set of CON sentences collected from 414 different online biomedical journal titles show that the SVM classifier yields a comparable result for the proposed a bag of unigrams input feature selectively extracted from a text of interest, compared to another bag of unigrams from the entire sentence. Moreover, we achieve a significant performance boost of the SVM with an input feature vector combining two types of statistical bag of unigrams and sentiment word lexicon.*

Keywords: Subjectivity classification, selective word statistics, Comment-on, support vector machine, MEDLINE

1 Introduction

MEDLINE® is the U.S. National Library of Medicine (NLM)'s premier online bibliographic database containing more than 26 million citations (including abstracts) from over 5,600 selected biomedical journals published in the United States and in other countries. Since the volume of biomedical literature is continually and rapidly growing, the number of journals indexed and the number of citations produced by NLM are also increasing dramatically; 130 journal titles are added each year on average, and nearly 806,000 citations were added to MEDLINE in 2015.

Users can access MEDLINE freely through NLM's PubMed [1], and open full-text articles through PubMed Central (PMC) [2]. These sites provide two conventional ways of navigating the enormous MEDLINE database to let users get the correct information: keyword-based searching and tracking citation links between an article and the external sources listed in the reference section. 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 (or PMC).

However, retrieving relevant articles or correct citation information from MEDLINE using these current methods could often be 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.

In order to improve the efficiency of searching, other highly popular and successful web-based scientific literature searching tools such as Google Scholar and CiteSeer [3] provide not only the aforementioned conventional searching methods, but also a citation count indicating how often a given article is cited by other articles. Thus, users could quite easily search and find works having a high impact or contribution on a certain research topic. However, like PubMed, these search tools also do not provide the author's reason for citing a particular article or other source.

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 often reflects the author's subjective opinions or sentiments as supportive, contrastive, corrective, etc. toward the cited works. This could be an important clue for researchers seeking relevant previous work or approaches for a certain research purpose. In this paper, we present a machine learning-based classification method for distinguishing subjective (positive or negative) citation sentences from objective (factual) ones in the body text of a given online biomedical article as the first step to analyze the citing author's sentiments toward the cited external sources.

A support vector machine (SVM) with a radial basis kernel function (RBF) is employed as our classifier. Input feature vectors for the SVM are created based on word-level statistics selectively calculated from a region of interest within a sentence that is found to have an actual expression of an author's opinion or sentiment. A lexical linguistic cue is also introduced for further improvement of the classification performance. We evaluated the performance of the SVM classifier in terms of accuracy, precision, recall, and F-measure rates for classifying the citation sentences containing an author's subjective opinions, specifically toward Comment-on (CON) articles. 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.

2 Subjectivity/Sentiment analysis

Sentiment analysis is an active research topic in the field of information retrieval (IR) and natural language processing (NLP), and mainly deals with opinions in documents or sentences which imply positive or negative sentiment [4]. Especially, there is a rapidly growing interest in automatically classifying author's sentiments expressed within citation sentences in scientific literature to provide effective tools for researchers who are seeking relevant previous works or approaches for a certain research purpose. Owing to a wide range of linguistic expressions and writing styles, identifying the citing author's sentiments toward the cited external sources expressed within citation sentences is still challenging.

Generally, automated classification of the author's sentiments in citation sentences can be considered a three-class classification problem, and would be solved through two separate and sequential classification tasks. The first task is to determine whether a given citation sentence contains an author's subjective sentiment (positive or negative) or not (objective). This task is usually called "subjectivity classification" [5]. The latter task is then to identify the polarity of sentiments expressed in those subjective citation sentences.

Approaches in most previous works on such subjectivity or sentiment classification can be divided into two categories: 1) Machine learning-based methods such as convolutional neural network (CNN) and SVM [6] [7], and 2) rule-based methods using sentiment lexicon [8]. The resulting analysis schemes have now begun to be employed in other areas, such as citation-based text summarization [9], bibliometrics [10], and social media monitoring [11].

3 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.

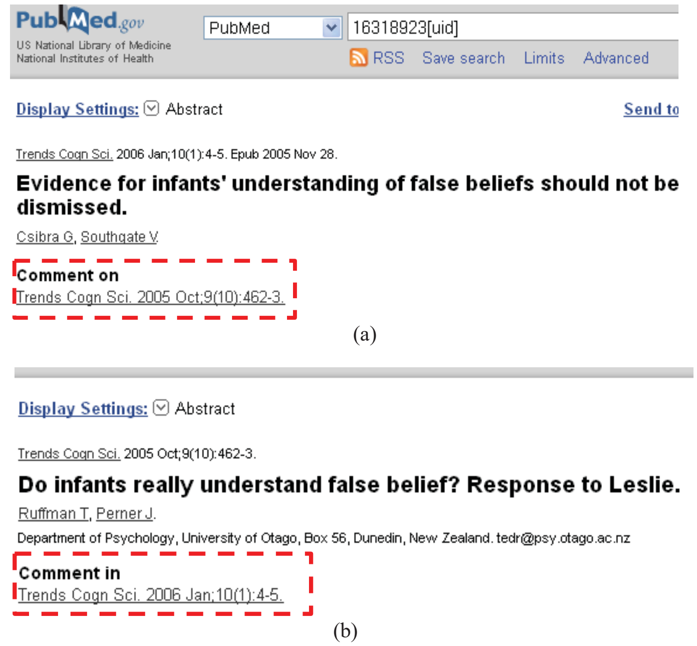


Fig. 1. (a) "Comment on" and (b) "Comment in" citations in MEDLINE

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.

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 sentiment: positive, negative, and neutral. Here, "neutral" represents the citing author's objective description of the cited work (neither positive nor negative). Our aim in this study is to automatically classify CON sentences having positive or negative citation sentiment into the "Subjective" class and neutral CON sentences into the "Objective" class, respectively. Typical examples of CON sentences in each category of author's citation sentiment are shown in Table 1.

Such CON sentences can be extracted through three preprocessing 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, and 3) identification of CON sentences from a set of citation sentences. We accomplished these preprocessing steps using machine-learning based methods developed in our previous studies [12].

Table 1. Examples of citation sentiments in CON sentence

Sentiments	CON sentences
Positive	We congratulate Krinsley and Jones (1) (2) for their impressive improvement in patient survival, which was attributed to tight glucose regulation.
	An ingenious method described by Yoon et al. (3) in this issue of PNAS now provides a powerful tool that can yield unprecedented information on membrane fusion at the single-vesicle level.
	Fleming et al [1] are to be commended for the excellent technical presentation of portal vein reconstruction using clear art work and intraoperative photographs.
Negative	The findings recently published by Schüz et al. (1), similar to all of the Interphone Study results published to date, have several serious problems.
	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.
	We disagree with Luty et al's suggestion that burpenorphine should replace methadone. [1]
Neutral	In 2003 Bardiau et al. (1) described the process of implementation of a nurse-based Acute Pain Service (APS) in a general hospital.
	In their Journal of Neuroscience article, Aziz-Zadeh et al. (2006) used functional magnetic resonance imaging to determine the extent to which the human mirror neuron system is lateralized.
	In the article by Najarian et al, (1) the authors evaluated the age-adjusted risk of stroke and population-attributable risk associated with either metabolic syndrome or type 2 diabetes mellitus in a cohort of 2097 adult subjects.

4 Proposed method

In this paper, we propose an automated method for citation subjectivity classification using an SVM-based text categorization technique and input feature vectors based on word-level statistical and lexical linguistic cues. Here, we focus on CON sentences in CIN articles first, but the proposed classification scheme could be easily extended to identifying author's subjective citations in other general online or offline biomedical articles.

4.1 Segmentation of text region of interest

Semantically, a CON citation sentence can be divided roughly into two parts: one for the citing author's opinion toward the cited external source and the other for a concise description of the cited work, as can be seen from the examples in Table 3. The latter (as shown as bold text in Table 3) actually has few or no contributions in correctly classifying the citing author's sentiment. Rather, it may cause the degradation of the classifier's discrimination performance.

Semantic analysis to locate a text region of interest containing an actual expression of an author's opinion or sentiment from a sentence is usually a complicated preprocessing procedure that requires a deep understanding of the grammatical structure of the sentence. Thus, most subjectivity/sentiment classification studies extract input features simply from the entire sentence. In order to overcome the problem of understanding the grammatical sentence structure, we use the "Stanford typed dependencies representation" [13] that was designed to provide a simple description of the grammatical relationships between words in a sentence.

Basically, the Stanford dependencies (SD) are triplets: governor (or head), its dependent, and the name of their grammatical relation. As can be seen in Fig. 2, such word dependencies in a given sentence map straightforwardly onto a directed graph representation in which words in the sentence are nodes and grammatical relations are edge labels in the graph. The main verb of a sentence is usually considered the "root" node in the SD graph. Generally, the main verb in citation sentences is found to belong to the abovementioned text region of interest describing the citing author's opinion.

Table 2. Examples of CON sentence semantically divided into two parts

We have read with great interest the original article by Bartels et al. (2004) in which **they showed that ventilation did not significantly affect spectral analysis of heart rate variability**.
Hooper's conclusions that **omega 3 fats have no effect on total mortality, combined cardiovascular events, or cancer** are somewhat misleading [1].
In this article, our French colleagues [1] have nicely demonstrated that **there is no advantage to delaying surgical closure after pharmacologic failure in premature infants in whom the PDA is documented to be hemodynamically significant.**

Based on this observation, our idea is to first analyze such grammatical relationship between words in a given sentence, and then to select words that are directly connected ("primary connection") to the "root" word. We also keep track of their dependent ("secondary connection") adjective/adverb words. As an example, in Fig. 2, the SD representation points to "read" as the root of the sentence. This root word is found to have a primary connection with the three words: "We", "interest", and "case". By tracking all adjectives and adverbs connected to these three words, we can also find the words of "secondary connection" (here, "intriguing" and "very"). We can see that all these "root", "primary connection", and "secondary connections" words are semantically linked for expressing the citing author's sentiment toward the cited work. We segment such text region of interest from the training set of CON sentences, and from which we calculate word-level statistics to build a bag of unigrams input feature vector for the SVM classifier.

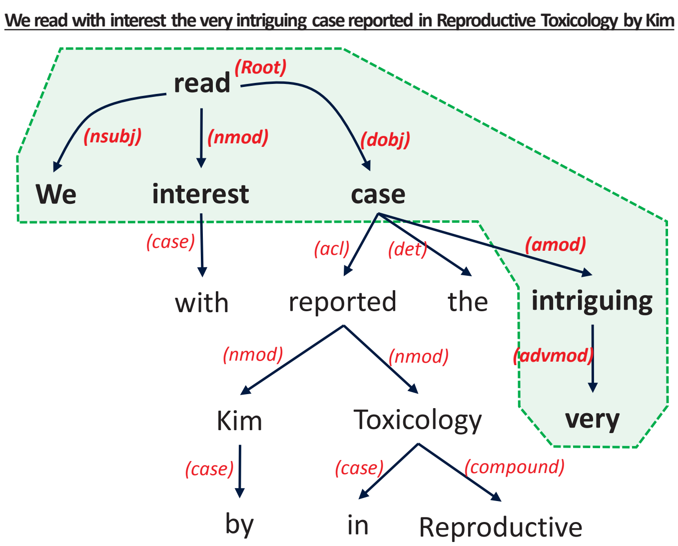


Fig. 2. Graphical representation of the Stanford dependencies for a given sentence.

4.2 Feature extraction

We employ two types of features to build an input feature vector for the SVM classifier: 1) bag of unigrams based on word statistics representing how differently a word is distributed in subjective and objective CON sentences and 2) sentiment lexicon consisting of positive or negative meaning of words. These features were experimentally found to be effective to separate subjective and objective CON sentences.

4.2.1 Bag of unigrams

Using words as input features requires a very high dimensional feature space (10,149 dimensions 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), χ^2 statistics, and simplified χ^2 ($s\chi^2$) statistics [14]. 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 χ^2 [15].

In our task, $s\chi^2$ of word t_k for CON sentences in the "Subjective" opinion class (class c_0) and those in the "Objective" class (class c_1) can be defined as follows:

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

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:

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

Accordingly, the more differently a word is distributed in "Subjective" and "Objective" opinion classes, the higher its $s\chi^2_{max}(t_k)$. Words are sorted according to their $s\chi^2_{max}$, and a word dictionary that is created by selecting words having the highest $s\chi^2_{max}$ scores is then considered a bag of words (unigrams) feature. Table 3 shows lists of the top 20 unigram words scoring the highest $s\chi^2_{max}$. 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. In the experiments, we built two types of such bag of unigrams: one (we call "selective unigrams") based on $s\chi^2_{max}$ selectively calculated from the aforementioned text region of interest within a CON sentence and the other from the entire CON sentence in the training set, and compared their classification performance.

Table 3. List of top 20 unigram words scoring the highest $s\chi^2_{max}$

Important, read, interest, article, issue, interesting, patients, provide, our, great, journal, new, further, provides, excellent, show, regarding, compared, paper, novel

4.2.2 Sentiment lexicon

Sentiment words indicate words in a language that have a positive or negative meaning. Table 4 shows several examples of positive and negative meanings of words. Most sentiment words are adjectives or adverbs, nouns (e.g., "pride" and "drawback") and verbs (e.g., "commend" and "disagree"), however, can also be used to express sentiments. Collectively, such sentiment words are called sentiment lexicon.

Such sentiment lexicon could not solely serve as an input feature for the subjectivity or sentiment analysis because identifying the author's real sentiment from a citation sentence is a much more complicated task. A citation sentence containing sentiment words may not express any subjective opinions. Conversely, a subjective sentence may not have any sentiment words. Moreover, words' sentiments are often heavily application domain or context dependent. For example, many biomedical terms or words (e.g., "cancer", "abuse", "disorder", etc.) frequently used in biomedical literature are labeled as negative in most of publicly available sentiment lexicons.

Therefore, sentiment lexicon has been used not alone but combined with other features in many sentiment or subjectivity classification studies [16][17]. We also combined this feature with the aforementioned two other features and evaluated how much it improves the overall performance of the SVM classifier in classifying citation subjectivity in CON sentences. We employed the sentiment lexicon constructed by Hu and Liu [18] which is publicly available and contains about 6,800 sentiment words in English. In this study, sentiment lexicon information is converted to a 2-bit binary vector of which each bit represents the existence of positive and negative words in a given sentence, respectively.

Table 4. Examples of positive and negative sentiment words

Positive	'excellent', 'wonderful, 'great', 'promising', 'fascinating', 'nicely', 'timely', 'elegantly', 'pride', 'state-of-the-art', 'contribution', 'congratulate', 'commend', 'appreciate', 'welcome'
Negative	'bad', 'wrong', 'incorrect', 'problematic', 'controversial', 'irrelevant', 'misleading', 'gravely', 'falsely', 'unfortunately ', 'drawback', 'error', 'weakness', 'conflict', 'dismay', 'disagree'

5 Classification experiments

5.1 SVM classifier and dataset

We employed an SVM [19] 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 [20].

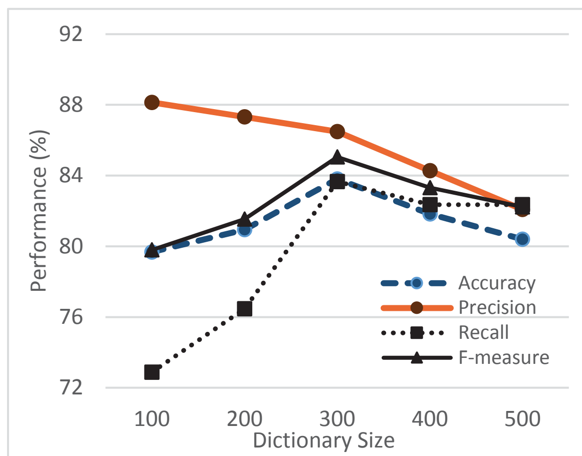
$$K(x_i,\, x_j) = exp(-\gamma \|x_i - x_j\|^2) \tag{3}$$

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 journals 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%).

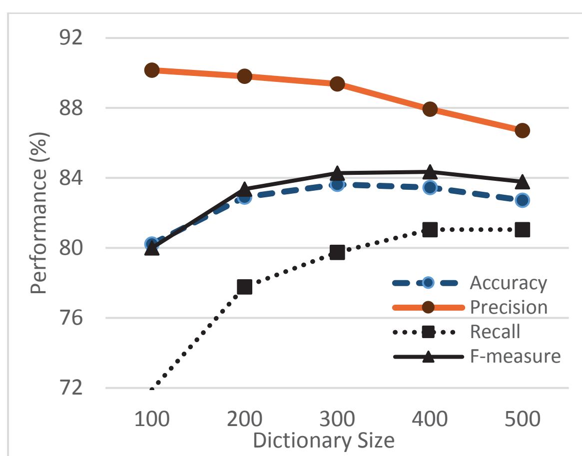
The collected CON sentences are then divided into two classes ("Subjective" and "Objective") according to the author's citation sentiment expressed within these CON sentences through a manual annotation process. Among these, 2,109 CON sentences consisting of 1,139 in the "Subjective" class and 970 in the "Objective" class were randomly selected to train the SVM. The statistics ($s\chi^2_{max}$) of words in the CON sentences are also estimated from this training set. The remaining 556 sentences (306 from the "Subjective" class + 250 from the "Objective" class) were used as a test set to evaluate the performance of the SVM.

5.2 Experimental results

In experiments, we first investigated the influence of word reduction in the abovementioned bag of words features to discover the best-performing word dictionary size. Figure 3 (a) 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 unigram words feature. As mentioned earlier, words in the dictionary are selected according to their corresponding $s\chi^2_{max}$ scores that reflect the difference of their distributions between the "Subjective" and "Objective" classes. We can see that our SVM classifier performs best overall when the size of the word dictionary is 300. Through a similar evaluation process, we also ascertained that the best-performing word dictionary size for another bag of unigrams ("selective unigrams") based on $s\chi^2_{max}$ derived from the region of interest within each sentence in the training set is 300, as can been seen in Fig. 3 (b). Although the SVM classifier yields slightly better Recall and F-measure rates when the dictionary size is 400, we choose the dictionary consisting of 300 words, considering the computational cost.



(a)



(b)

Fig. 3. SVM performance against different word dictionary sizes in the bag of (a) unigrams and (b) selective unigrams.

Next, we compared the performance of the following four types of input feature vectors for the SVM classifier.

1) *Input feature 1*: bag of selective unigrams

2) *Input feature 2*: bag of unigrams

3) *Input feature 3*: input feature 1 + input feature 2

4) *Input feature 4*: input feature 3 + sentiment lexicon

From the classification results included in Table 5, the overall performance of "*input feature 1*" extracted only from a text region of interest within a sentence is found to be comparable to that of "*input feature 2*" from the entire sentence. More importantly, the SVM classifiers yields a significantly better performance overall for "*input feature 3*", the combination of these two types of bag of unigrams. Thus it is analyzed that each bag of unigrams input feature can compensate substantial errors in the other. Further improvement of performance is also achieved by additionally introducing sentiment lexicon information (*input feature 4*).

Table 5. The subjectivity classification performance of SVM classifier for different input features

Input features	Accuracy	Recall	Precision	F-measure
Input feature 1	83.63	79.74	89.38	84.28
Input feature 2	83.81	83.66	86.49	85.05
Input feature 3	87.41	88.89	88.31	88.60
Input feature 4	**88.12**	**90.20**	**88.46**	**89.32**

Table 6 shows examples of false-negative (FN) and false-positive (FP) classification errors from the SVM. The first CON sentence in the FN error examples contains the word, "*groundbreaking*" suggesting a positive sentiment. This word can also be found in our sentiment lexicon but is missing in the word dictionary of the bag of unigrams input features, certainly due to the small size of our current training dataset. Thus we expect that this type of errors could be fixed by collecting more CON sentences and increasing the size of the training dataset. The SVM input features adopted in our study are all "word-level" features. Thus the SVM classifier often encounters some difficulties in dealing with the citation sentences which only use subjective phrases to convey author's sentiment. As an example, the second CON sentence of FN error in Table 6 is misclassified as "Objective" even though it apparently has a positive meaning phrase ("*breaks new ground*").

The FP error example shown in Table 6 has a strongly positive word, "*excellent*". However, this word is used not for representing the citing author's sentiment but for just describing the cited work. As previously mentioned, input features based on a bag of unigrams and sentiment lexicon consider the entire sentence to identify an author's sentiment. As a result, the SVM using these input features made a misclassification error for this CON sentence. On contrary, the proposed bag of selective unigrams is extracted only from the text of interest being expected to have an author's opinion or sentiment toward the cited work. In our experiments, the SVM using the bag of selective unigrams input feature is found to successfully classify this CON sentence into the "Objective" class.

6 Future work

As future work, we first plan to employ a deep learning technique such as CNN as another classifier and compare its performance with our current SVM classifier. We will verify that our multi-channel approach considering both the entire sentence and only text region of interest can also boost the deep learning performance. In this study, through a series of experiments and error analysis, our ground-truth training dataset was found not big enough to reliably calculate word-level statistics employed to create the bag of words input features for the SVM classifier and for future deep learning task. Therefore, we are next 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.

Finally, based on our achievement of this subjective analysis study, we will move on to the next stage of our task, developing a reliable method to identify the polarity of the citing author's sentiments expressed in the subjective citation sentences.

Table 6. Error examples showing false negative and false positive errors

Error types	CON sentences
False-Negative	This is exactly what de Jonge et al. [4] achieved with their *groundbreaking* investigation presented in this issue of the Journal.
	The report published by Menke et al [7] in this issue of Circulation *breaks new ground* by extending the dose-effect relation to considerably lower blood lead concentrations than reported in previous studies. [3-6]
False-Positive	Naguib et al. [1] examined the dose of succinylcholine required for *excellent* tracheal intubating conditions.

7 Conclusions

Identifying author's sentiment expressed within citation sentences in scientific literature could be an important tool for researchers seeking relevant previous work or approaches for a certain research purpose. In this study, we have presented a machine learning-based automated classification method for distinguishing subjective citation sentences from objective ones in the body text of an online biomedical article, as the first step to identify the citing author's sentiments toward the cited work.

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 for classifying the citation sentences containing an author's subjective opinions 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. We have introduced a bag of selective unigrams based on word statistics calculated from a text region of interest within a sentence, instead of from the entire sentence, as an input feature for the SVM classifier. This text region of interest reflects the Stanford dependencies (SD)-based grammatical relationships between words in a sentence, and is found to have an actual expression of an author's opinion or sentiment.

Through a series of experiments on a set of CON sentences collected from 414 different online biomedical journal titles, we see that the overall performance of the proposed bag of selective unigrams extracted only from a text region of interest is comparable to that of another bag of unigrams from the entire sentence. Moreover, the combination of these two different types of bag of unigram input features significantly boosts the performance of the SVM classifier, indicating that each bag of words feature can compensate substantial errors in the other. Further improvement is also achieved by additionally employing sentiment lexicon information. Future research is required to deal with errors resulting from our current approach, which would include 1) increasing the size of the ground-truth dataset and 2) employing and testing deep learning techniques.

Acknowledgment

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

8 References

[1] http://www.ncbi.nlm.nih.gov/pubmed

[2] http://www.ncbi.nlm.nih.gov/pmc/

[3] 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.

[4] B. Liu, *Sentiment analysis and opinion mining*, Morgan and Claypool, 2012.

[5] J.M. Wiebe, R.F. Bruce, and T.P. O'Hara, "Development and use of a gold-standard data set for subjectivity classifications," *Proc. the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics (ACL' 99)*, pp. 246-253, Maryland, June 1999.

[6] Y. Kim, "Convolutional neural networks for sentence classification," *Proc. Conf. Empirical methods in natural language processing (EMNLP'14)*, pp.1746-1751, Doha, Qatar, Oct. 2014.

[7] A. Athar, "Sentiment analysis of citations using sentence-structure-based features," *Proc. ACL-HLT 2011*, pp. 81-87, Portland, June 2011.

[8] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267-307, June 2011.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] M. de Marneffe and C.D. Manning, "Stanford typed dependencies manual," (http://nlp.stanford.edu/downloads/lex-parser.shtml), 2008.

[14] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1-47, 2002.

[15] 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.

[16] H. Yu and V. Hatzivassiloglou, "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences," *Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP2003)*, pp. 129-136, 2003.

[17] M. Hu and B. Liu, "Mining and summarizing customer reviews," *Proc. ACM SIGKDD int'l Conf. Knowledge Discovery and Data Mining (KDD 2004)*, pp. 168-177, Seattle, Aug. 22-25, 2004.

[18] M. Hu and B. Liu, Sentiment lexicon, (http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html)

[19] V. Vapnik, *The nature of statistical learning theory*, New York: Springer-Verlag, 1995.

[20] C.C. Chang and C.J. Lin, "LIBSVM: a library for support vector machines," 2001. [*http://www.csie.ntu.edu.tw/ ~cjlin/libsvm*]