An Application of Aspect-Based Sentiment Analysis on Teaching Evaluation

Wen Qu[1] and Zhiyong Zhang[1]

University of Notre Dame, Notre Dame IN 46556, USA
wqu@nd.edu
zzhang4@nd.edu

Abstract. With the rapid development of new techniques, text mining has become explosively popular in the recent two decades. Various techniques and methods have been developed to manage and analyze text data to exploit the information underlying the text. Among them, the aspect-based sentiment analysis (ABSA), which is a research field that studies people's opinion, sentiment toward attributions or aspects of individual entities, has attracted researchers in both industry and academia. ABSA first extracts the relevant aspects of a specific entity and then determines the sentiment for each aspect. To our knowledge, there is no ready-to-use `R` packages or functions for ABSA. In this study, we conduct a brief review of ABSA and apply it to a teaching evaluation study. We also illustrate how to conduct ABSA using `R`.

Keywords: Aspect-based sentiment analysis · Teaching evaluation · Text data.

1 Introduction

In daily life, when we need to make a decision, we often consider others' opinions, such as purchasing a product, searching for a restaurant, and choosing a doctor. With the rapid growth of the Internet, almost every aspect of life has been changed dramatically. How to make decisions is also changing. Instead of asking family members and friends for advice, people tend to use social media for help nowadays. For example, when someone wants to buy a new computer, he or she may search for comments on the websites of online retailers. One can obtain valuable information from others who have a similar experience by simply glancing over a few number of comments.

However, when there are many comments and reviews, it can be difficult for people to summarize information quickly. How to extract useful information from the texts is difficult but critical, especially in the digital age. Text mining, therefore, becomes very popular in the recent two decades with the growth of social media (Liu, 2015). Sentiment analysis is one of the various techniques and methods for managing and analyzing text data to exploit the underlining information.

Text mining is a field that analyzes people's opinions and sentiments towards certain entities, and their aspects (or attributes) expressed in the text (Liu & Zhang, 2012). It got popular after 2000 and was originated from computer science, but its applications have spread to business, management, sociology, political science, and literature. Moreover, its research has been mainly carried out at three levels: document, sentence, and aspect. In this study, we limit the scope to the last one. Comparing to both document and sentence level analyses, the aspect-based sentiment analysis (ABSA) can return the sentiment of aspects or the attributes of an entity instead of the overall polarity of the entity. In other words, we get to know the specific target of opinion with ABSA.

For an ABSA, with a particular entity, there are two main tasks: first, to extract the aspects from the text that needs to be evaluated, and second, to classify the sentiment for each aspect. In the literature, the aspect extraction task involves expression extraction and grouping. In Liu's (Liu, 2015) book, he gave a comprehensive review of the well-known methods and their applications. Due to the page limitation, we only review the related methods that have been used in our study. The most straightforward approach to extract aspect expression is through the frequency of nouns or noun phrases (Blair-Goldensohn et al., 2008; Hu & Liu, 2004). This method is simple but effective because the nouns are often used to describe the aspects, and the vocabulary people use tends to be similar when they talk about different aspects under the same entity. By applying some rules, researchers can keep specific frequent nouns as aspect expressions. With the expressions, unsupervised methods (i.e., using dictionaries to find synonymous expressions or expert labeling) are applicable for grouping the expressions when the entity requires less specific knowledge (e.g., general product review and teaching evaluation). The second task, sentiment classification, can be accomplished by supervised learning methods (Jiang, Yu, Zhou, Liu, & Zhao, 2011) or unsupervised lexicon-based methods (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011).

The ABSA is a domain-sensitive method; that is, a well-developed method for some domains may not be suitable for a different domain. To our knowledge, there is no method available in the literature with regards to teaching evaluation. Depend on the type of audience, the evaluation of teaching might have different purposes (Ory, 2000), including gathering feedback for teaching improvement, collecting data for personnel decision-making, or providing options for course selection. In the digital age, the form of teaching evaluation switched from the paper to the online version with both quantitative data and qualitative comments. This change makes it possible to conduct a more comprehensive mixed analysis like the longitudinal studies or text mining with hundreds or thousands of evaluation records. Therefore, in the rest of the paper, we apply ABSA to a teaching evaluation study and illustrate how to conduct the analysis using an R function we developed.

2 Application

In this section, we show how to apply the ABSA to the text comments on teaching evaluation. We first describe the data used in this study. Then, we propose a two-stage procedure to conduct ABSA. After that, we illustrate the procedure through the real data.

2.1 Data

The data were crawled from the *ratemyprofessor.com* website by conforming to their crawling rules. The crawled dataset contains 1954 records of teaching evaluation from students for 50 college professors in the United States. Each record is an evaluation from one student on one professor. The sample size is considerably large in ABSA studies.

The dataset includes five variables. The first variable is the identification number of a record. The second variable is the unique id of a professor. The third variable is the numerical rating of a professor with 1 to 5 indicating worst to best to the answer of a question — "How would you rate this professor as an instructor?" The fourth variable is the date of a comment. The time of the records ranges from the year 1999 to 2018. The last variable is the response to an open-ended question about overall evaluation of a professor. Every student who filled out the evaluation form was required to write a short comment here.

The histogram of the number of evaluations received by each professors is given in Figure 1. The number ranges from 8 to 98, and on average, each professor received 39 evaluations (the median is 33).

Two histograms of the rating scores are displayed in Figure 2. The left histogram is for the 1954 comments, and the right histogram is for the averaged rating score of each professor. With the grouping variable (here is the 'professor id'), the distribution of the rating scores show a bimodal pattern to a negative skewed pattern.

The comments are in the free text format with a maximum of 350 characters for each. An example comment is given below.

> *Great teacher, really know his stuff. Also use a TON of example to try to explain everything well. Will give example of crime and let the class talk it out. You are free to ask question and he will answer very well. Also put his note on blackboard/hand them out in class, great to study with. I would suggest him.*

This comment clearly expressed a general positive sentiment towards the professor under evaluation. Our goal is to understand the sentiment of all 1,954 comments in the dataset.

2.2 Procedure

The procedure for evaluating teaching based on the text comments using ABSA involves two sequential stages. First, we discuss a combined method with word frequency, lexicon, and human labeling to extract aspects. Second, we develop

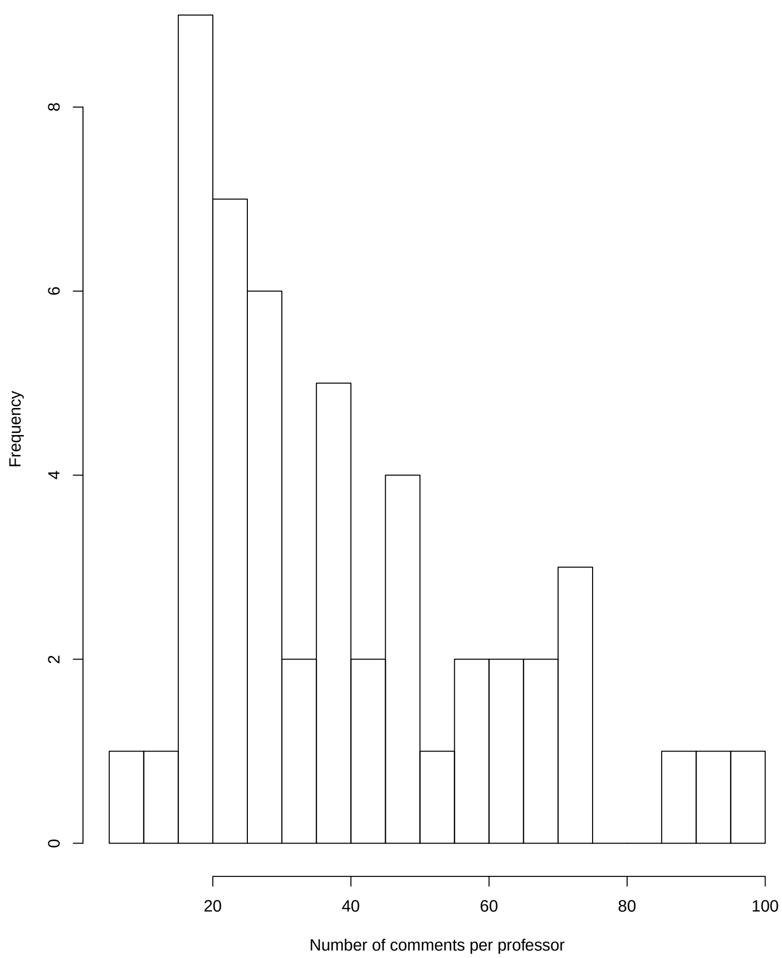

Fig. 1: Histogram of number of evaluations received by each professor

an unsupervised scoring method for aspects' sentiment. The details are provided as follows.

Aspect Extraction and Grouping. The first task of ABSA is always to extract aspect expressions and then group them into aspect categories. We illustrate the procedure using the teaching evaluation data. We began by selecting all the nouns from all the comments, which led to a total of $n = 1511$ words.

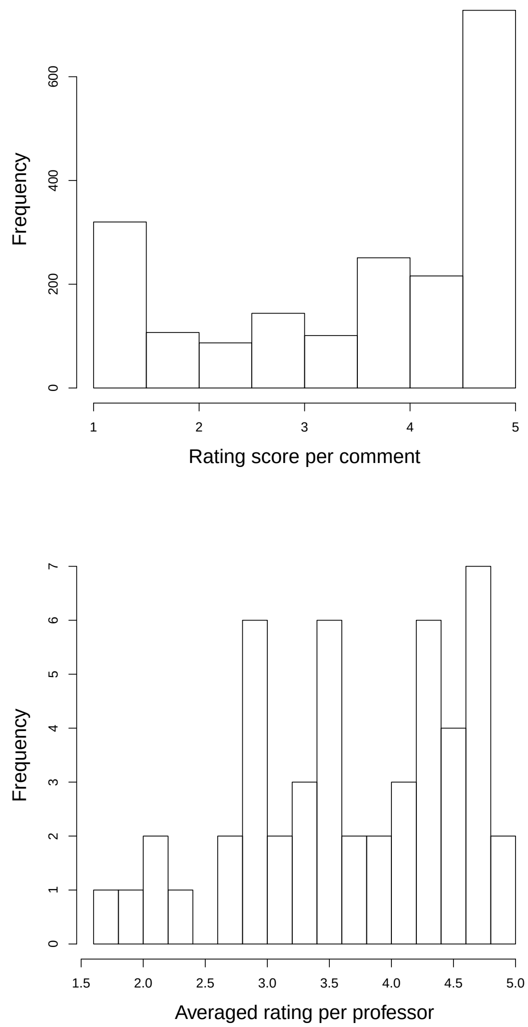

Fig. 2: Histograms of rating scores

We then applied the following four constraints to obtain a shorter but more measurable and meaningful list:

(1) removing the stop words which are commonly used English words not having important meanings;
(2) removing nonsense nouns which are not meaningful in the context of the specific entity (in this example, the entity is teaching evaluation), such as 'bit', 'ton';

(3) adding customer-define words which are not in the previous word list, such as ‘his’,‘her’;
(4) setting the frequency threshold to remove less used words ($freq > 9$).

The remaining words were the candidates of the aspect expressions ($n = 189$) such as ‘professor’, ‘homework’, and ‘quiz’.

Based on our knowledge on teaching evaluation, we decided to have five aspect categories: (A1) class in general; (A2) teacher in general; (A3) exam; (A4) grade; (A5) workload and class activity. The *class in general* aspect includes the expressions indicating class environment, course structure, materials, and the other general perspectives of a class. The *teacher in general* aspect describes the personality, teaching style, and research ability of a professor. The *exam*, *grade*, and *workload and class activity* aspects measure the other teaching evaluation related facets.

After having the expressions candidates and deciding the categories, we first labeled some expressions candidates as seed set and got their synonymous from the unlabeled set. We then manually labeled the rest based on the general knowledge about teaching evaluation. In total, we extracted 110 aspect expressions under five categories. Table 1 shows the aspect categories with their corresponding expressions.

Aspect Scoring. In the current literature and applications, the majority of researchers focus on the orientation of the sentiment. To facilitate future analysis with other numerical variables (e.g., rating scores), we focus on the intensity of the sentiment—getting the sentiment scores for each aspect.

We used a four-step algorithm to score the aspects. First, for each comment, we scored the sentiment expression using the AFINN dictionary (a word list with discrete ratings of 2477 sentiment expressions). Concurrently, we added the aspect category label to each aspect expression (e.g., an aspect expression ‘teacher’ was labeled as its category ‘A2’). Second, we applied the negation shifter (such as ‘not’, ‘no’, ‘neither’, ‘nor’) for sentiment ratings to change the direction of the negation sentiment expressions. Third, with aspect expressions and scored sentiment expressions, we applied the syntactic dependency rules to link them and assigned the score of a sentiment expression to its corresponding aspect expression. The syntactic dependency rules define the grammar relation between two words in a sentence with one word being the root and the other being the dependent. According to the Stanford Dependencies manual (de Marneffe & Manning, 2008) and our data, we selected 11 rules to quantify the dependency relationship between the aspect expressions and the sentiment expressions. Fourth, we aggregated the sentiment scores per aspect of each comment, and then got the averaged scores for each professor.

Example We use a simple example with only one sentence to demonstrate the procedure of the aspect scoring algorithm. The sentence is *”I love the prof, but the exam is not easy and long”*. In the first step, we identified the aspect expressions with their categories (*prof*→A2,*exam*→A3) and scored the sentiment

Table 1: Aspect categories and expressions

Aspect Category	**Size**	**Aspect Expressions**
A1 (Class in general)	N1=52	class lecture mathematics hour book material text note stuff speech subject topic guide tetbook chapter history powerpoint office chemistry econ lab art concept forum information syllabus music email video skill handout participate business instruction term expectation finance literature science approach classroom explain requirement yoga calculus line require movie schedule animal detail english
A2 (Teacher in general)	N2=30	professor teacher guy teaching person prof dr teach instructor experience sytle humor accent knowledge research writer prof. he she his her him himself herself he's she's
A3 (Exam)	N3=9	exam test essay final midterm paper quizzes quiz quizze
A4 (Grade)	N4=10	grade credit attendance grading attention level pass score gpa bonus
A5 (Workload & class activity)	N5=9	homework assignment time project reading writing practice discussion presentation

expressions (**love**→3,**easy**→1). Second, with the negation shifter, we changed the direction of the negative sentiments (**not easy**→-1). Third, we applied the syntactic dependency rules to link the sentiment expression with its aspect expression (*prof* $\overset{\text{obj}}{\Longrightarrow}$**love**, *exam* $\overset{\text{nsubj}}{\Longrightarrow}$**not easy**). Specifically, the *obj* rule indicates that the aspect expression *prof* is the direct object of sentiment expression **love**. The *nsubj* rule shows the aspect expression *exam* is the subject of the sentiment expression **easy** which has been changed the direction by the negation shifter ('not') in the previous step. Figure 3 shows the syntactic dependency relationship of the sentence.

In the last step, we aggregated the scores for each aspect category. In summary, in this example, only two aspects were involved in the sentence, and scores for them were 3 (A2: teacher aspect) and -1 (A3: exam aspect).

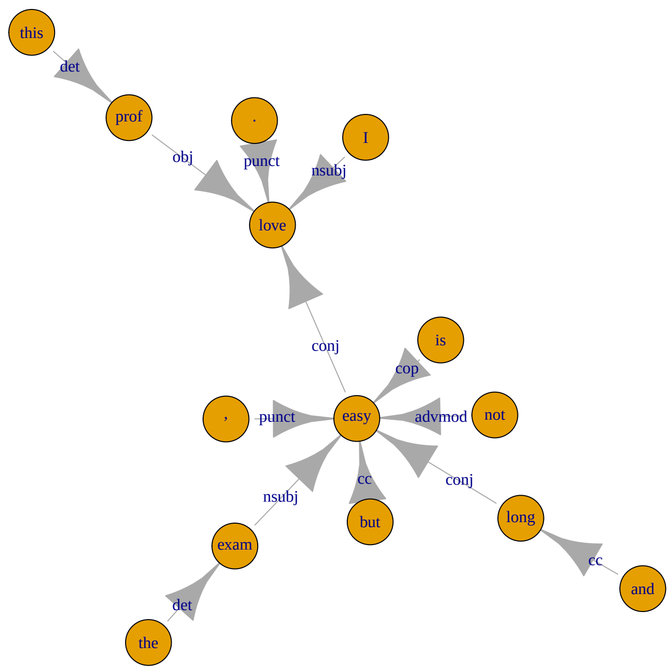

Fig. 3: The syntactic dependency relationship of one sentence (*I love the prof, but the exam is not easy and long.*)

2.3 Results

In the current study, we analyzed almost two thousands of comments to get the averaged aspect scores for each professor. Table 2 shows the scores of the 5 aspects, number of comments, and the averaged rating scores for the 50 professors. A positive/negative aspect score means the professor received an overall positive/negative opinion in that aspect among all comments he or she received. Because all the aspect scores were averaged, score 0 means the positive and negative opinions were canceled out. NA means the professor got no comment containing any expressions of that aspect. It usually happened when there were not enough comments like the professor with ID 3, who only had eight comments. If one aspect has a large amount of NA values, the researchers may change that aspect category, like adding more aspect expressions, combining with another category, or deleting the current one.

Different combinations of the aspect sentiments indicate distinctive teaching types. For example, professor 12 had 36 comments, and students positively evaluated his/her class structure and teaching style and course workload, but had a negative evaluation about the exam. Professor 40 had a similar amount of comments, but the evaluation was opposite. He/she received overall negative evaluations on the class, exam, grading, and workload aspects, but a positive opinion on the teaching aspect.

The histograms and the correlations of the five aspects are shown in Figure 4. The class in general (A1) and teacher in general (A2) aspects are moderately positively correlated ($r = 0.49, p < 0.001$). Since the first and the second aspects contain most of the aspect expressions ($n_1 = 52, n_2 = 30$) and are more related to the teaching evaluation covered in the written comments, most of the students have similar opinions on these two aspects. For example, if a student thinks his or her professor is good at teaching (A2), very likely, he or she would also have a positive opinion of the class (A1). This is because a teacher's behavior usually profoundly affects the quality of the class in general.

The other aspects have weakly or none correlations, indicating that students rarely comment on those aspects together. Another explanation could be because only a few expressions were involved in the last three aspects. Thus, those aspects might be hard to be found in the comments. Future studies therefore need to consider adding more expressions in those aspects or redesigning the aspect categories.

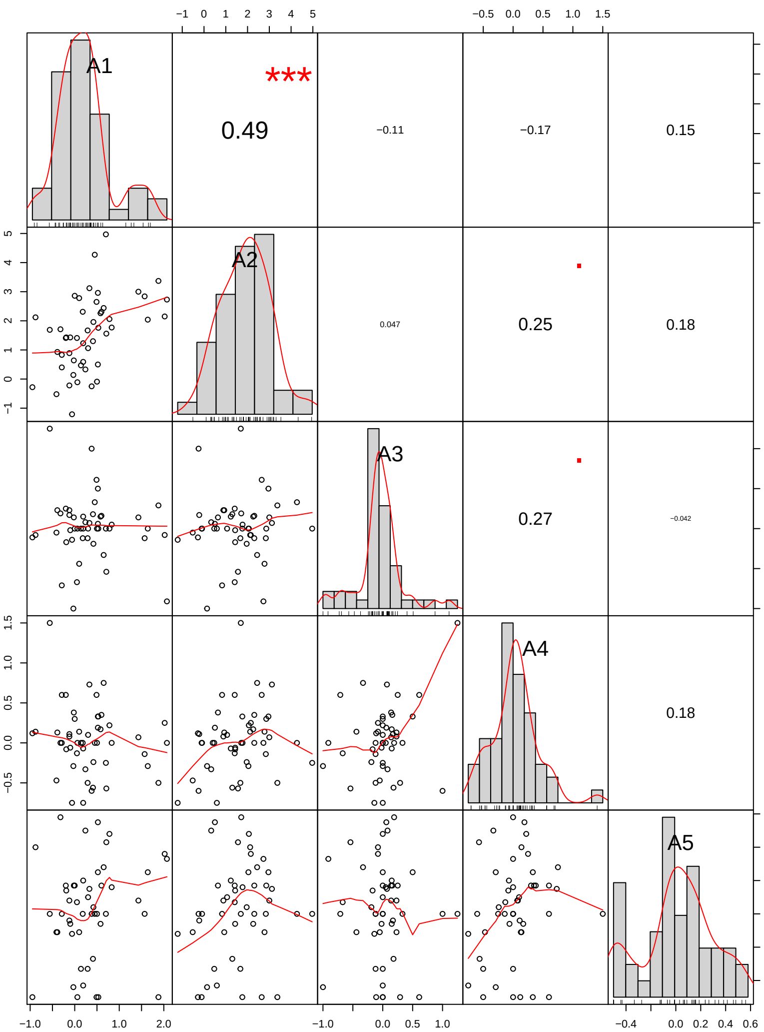

Fig. 4: Histograms and correlations of five aspects (* $p \leq 0.05$, ** $p \leq 0.01$, ***$p \leq 0.001$)

Table 2: Aspect scores of 50 professors (Ncom: number of comments)

ID	Ncom	Rating	A1(Class)	A2(Teacher)	A3(Exam)	A4(Grading)	A5(Workload)
1	28	4.25	0.29	1.67	-0.12	-0.5	-0.33
2	50	3.78	-0.19	1.43	-0.17	-0.08	0.14
3	8	3.69	-0.29	0.4	NA	0	NA
4	47	4.03	0.65	2.44	-0.33	0.75	0.28
5	72	2.94	-0.39	0.93	0.23	0.13	-0.11
6	98	3.83	0.19	1.23	0.15	-0.07	0.2
7	18	4.22	-0.56	1.69	1.25	1.5	0
8	17	3.41	-0.2	1.41	0.25	0.6	0.17
9	73	4.36	0.78	2.06	0	0.22	0.48
10	60	4.42	0.1	2.78	-0.44	0.14	-0.11
11	17	2.24	0.38	-0.25	1	-0.6	0
12	36	4.65	2.07	2.73	-0.91	0	0.33
13	66	3.47	-0.1	1.43	-0.02	-0.06	-0.06
14	21	2.1	-0.95	-0.28	-0.11	0.12	-0.5
15	27	4.28	1.64	2.04	0	-0.29	0.25
16	22	4.82	0.45	4.27	0.33	0	0
17	18	3.03	0.14	0.47	0	0	-0.33
18	18	4.5	0.18	2.31	-0.12	0	0
19	35	3.51	-0.32	1.71	0.19	0	0.58
20	74	2.14	-0.12	-0.22	0.17	0.11	-0.04
21	20	3.38	0.71	1.56	-0.54	-0.57	0.43
22	44	4.68	0.49	2.65	0.61	0.6	-0.5
23	47	4.49	2.02	2.15	-0.08	0.25	0.36
24	18	3.36	-0.29	0.83	-0.71	0.6	0
25	60	3.84	0.42	1.96	-0.19	-0.24	0.04
26	53	2.94	-0.02	0.64	0.14	0.38	0.17
27	28	3.5	-0.12	0.89	0.23	0.08	0.08
28	36	3.47	0.52	0.5	0.06	0.19	0.55
29	18	4.44	0.53	1.76	0	0.33	-0.5
30	61	4.13	0.41	1.3	0.18	-0.56	-0.27
31	38	4.36	1.57	2.84	-0.12	-0.14	0
32	39	3.32	0.05	1.41	-0.67	-0.13	0.07
33	22	2.86	0.06	-0.11	0	0	-0.5
34	17	2.71	-0.88	2.12	-0.08	0.14	0.4
35	12	2.92	0.5	-0.09	0	0	0
36	94	4.08	0.6	2.31	0.16	0.35	0.17
37	31	4.85	0.7	4.97	0	-0.25	0
38	28	4.7	1.88	3.37	0.29	-0.5	-0.5
39	49	4.66	0.83	1.77	0.05	0	0.16
40	37	2.62	-0.03	0.14	-1	-0.29	-0.44
41	86	1.71	-0.41	-0.52	-0.05	-0.47	-0.11
42	67	3.6	0.58	2.26	0.15	0.17	-0.06
43	26	4.73	0.33	3.12	0.07	0.73	0.15
44	21	3	0.24	0.33	0.08	-0.33	0.5
45	21	1.81	-0.06	-1.21	-0.14	-0.75	-0.12
46	22	3.11	0.19	0.59	0	-0.75	-0.43
47	41	3	0.3	1.06	0	0.1	0.1
48	65	4.77	1.43	3	0.14	0.07	0.08
49	26	4.71	0.52	2.96	0.5	0.33	0.25
50	22	4.36	0	2.86	0	0.3	0.17

3 R Function

We also developed an R function to implement the aspect scoring algorithm, which is provided in the appendix. There are three arguments in the function:

- *comments*: A list of comments.
- *aspect.list*: A list of customized aspect categories with expressions.
- *cinf*: The additional information of the comments. If there is additional information (e.g., professor id, time), users can put them in a data frame and assign it to the cinf argument. The default value is NULL.

The output depends on whether there is additional information. If there is no additional information, the output would be the comments with the sentiment scores of each aspect category. An example of two comments with R input and output is given below.

```
> comment=c("Class is not easy but don't get discourage after the midterm. The final is not bad. He is a
      good prof, but the lecture is useless. Focus on the materials.", "I love this prof, but the exam is
      not easy and long." )
> example=aspect_scoring(comment,aspect_hl)
Joining, by = "word"
Joining, by = "word"
> example
  comment_id A.1 A.2 A.3
1          1  -3   3   3
2          2  NA   3  -1
```

From the result, we found that the first comment included three teaching evaluation aspects — class (A1), teacher (A2), and exam (A3). The sentiment scores were consistent with the comment, in which the student explained the class was not easy, but the exam and teacher were good. The second comment was from the example in section 2.2. Because the second one only covered two aspects, the aspect not present in the comment had the score as NA.

On the other hand, if there is additional grouping information, like our data (grouping variable: professor ID), the output would return a table of professors with averaged sentiment scores of each aspect category averaging over the corresponding comments (like in Table 2). Users can easily adjust this option by modifying the R function to fulfill their need to obtain the averaged aspect scores based on people, time, institute, and other grouping variables.

4 Conclusion

In this study, we developed an approach for teaching evaluation using ABSA. We generated an aspect list for this domain and provided an R function for aspect sentiment scoring. By showing the procedure step by step and providing the R function, we hope to shine a light on the ABSA in the teaching evaluation field. With the aspect sentiment scores, other data analysis can be conducted, e.g., a mixed analysis of both text data and quantitative data.

In the future, instead of using the existing sentiment dictionary such as AFINN, we will develop a domain-specific lexicon for better sentiment expression identification. Furthermore, we plan to gather information from both students and teachers about the teaching evaluation for a more precise aspect exploiting.

References

Blair-Goldensohn, S., Hannan, K., McDonald, R. T., Neylon, T., Reis, G. A., & Reynar, J. (2008). Building a sentiment summarizer for local service reviews..

de Marneffe, M.-C., & Manning, C. D. (2008). Stanford typed dependencies manual [Computer software manual]. Retrieved from https://nlp.stanford.edu/software/dependencies_manual.pdf

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth acm sigkdd international conference on knowledge discovery and data mining* (pp. 168–177). doi: https://doi.org/10.1145/1014052.1014073

Jiang, L., Yu, M., Zhou, M., Liu, X., & Zhao, T. (2011). Target-dependent twitter sentiment classification. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies - volume 1* (pp. 151–160).

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments and emotions.* Cambridge University Press.

Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415–463). Boston, MA: Springer US. doi: https://doi.org/10.1007/978-1-4614-3223-4_13

Ory, J. C. (2000). Teaching evaluation: Past, present, and future. *New Directions for Teaching and Learning*, *2000*(83), 13–18. doi: https://doi.org/10.1002/tl.8302

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011, June). Lexicon-based methods for sentiment analysis. *Comput. Linguist.*, *37*(2), 267–307. doi: https://doi.org/10.1162/COLI_a_00049

Appendix: R function for aspect scoring

```r
library(wordnet)
library(NLP)
library(coreNLP)
library(cleanNLP)
library(udpipe)
library(tidyverse)
library(tidytext)
library(dplyr)
library(lubridate)
library(reshape2)
library(lmerTest)
library(lexicon)

# Using for dependency parsing
udmodel_en = udpipe_load_model(file = "english-ewt-ud-2.3-181115.udpipe")

#input: a set of comments of teaching evaluation;a list of aspects
#output: 5 aspects scores
aspect_scoring = function(comments,aspect.list,cinf=NULL){

  comments = tolower(comments)
  size = length(comments)
  annotate = udpipe_annotate(udmodel_en,x=comments,doc_id=c(1:size))
  dataframe.annotate = as.data.frame(annotate)
  negation=c('none','not','never','neither','nobody','nowhere')
  #aspect class and sentiment score for corresponding word
  ap_rs = dataframe.annotate%>%
    select(doc_id,sentence_id,token_id,lemma,head_token_id,dep_rel)%>%
    filter(!dep_rel=='punct')%>%
```

```r
    mutate_at(c('doc_id','sentence_id','token_id','head_token_id'),funs(as.numeric))%>%
    rename(word=lemma)%>%
    left_join(get_sentiments('afinn'))  %>%
    left_join(aspect.list) %>%
    mutate(id=NULL,neg=if_else(word %in% negation,1,0),sentiment=NA)

  list.ap_rs=split(ap_rs,ap_rs[,1:2])
  list.ap_rs=list.ap_rs[sapply(list.ap_rs,function(x) dim(x)[1])>0]
  for (s in 1:length(list.ap_rs)){
    for (i in 1:dim(list.ap_rs[[s]])[1]){
      if (list.ap_rs[[s]][i,]$neg == 1 ){
        hid=list.ap_rs[[s]][i,]$head_token_id
        list.ap_rs[[s]][list.ap_rs[[s]]$token_id==hid,]$value = -list.ap_rs[[s]][list.ap_rs[[s]]$token_id==
              hid,]$value
      }
    }
  }
  get_aspect_sentiment=function(ls){#ls: list of annotation file
    dep_a2o=c('nsubj','obj','obl','nmod','conj','advcl','xcomp','amod','acl:relcl','advmod','acl','obl:tmod'
          ,'obl:npmod','iobj')
    for (s in 1:length(ls)){
      for (i in 1:dim(ls[[s]])[1]){
        if (!is.na(ls[[s]][i,]$Aspect) & ls[[s]][i,]$dep_rel %in% dep_a2o){
          hid=ls[[s]][i,]$head_token_id
          if(hid %in% ls[[s]]$token_id){
            ls[[s]][i,]$sentiment = ls[[s]][ls[[s]]$token_id==hid,]$value
          }
        }
        else if (!is.na(ls[[s]][i,]$value) & ls[[s]][i,]$head_token_id !=0){
          hid=ls[[s]][i,]$head_token_id
          if(hid %in% ls[[s]]$token_id){
            if (!is.na(ls[[s]][ls[[s]]$token_id==hid,]$Aspect) & is.na(ls[[s]][ls[[s]]$token_id==hid,]$
                  sentiment))
              ls[[s]][ls[[s]]$token_id==hid,]$sentiment = ls[[s]][i,]$value
          }
        }
      }
    }
    return(ls)
  }
  list.ap_rs.final=get_aspect_sentiment(list.ap_rs)
  data.ap_rs.final=as.data.frame(bind_rows(list.ap_rs.final))
  doc_aspect_sentiment=data.ap_rs.final%>%group_by(doc_id,Aspect)%>%summarise(score=sum(sentiment,na.rm=T))
        %>%filter(!is.na(Aspect))%>%rename(A=score)
  doc_aspect_sentiment_wide=reshape(data.frame(doc_aspect_sentiment),timevar='Aspect',idvar='doc_id',
        direction='wide')
  doc_aspect_sentiment_wide=doc_aspect_sentiment_wide%>%rename(comment_id=doc_id)
  #the setting only for current study with 5 aspect,users can change it accordingly
  if (is.null(cinf))
    rst=doc_aspect_sentiment_wide
  else{
    final.com.abs=cinf%>%left_join(doc_aspect_sentiment_wide)
    final.pf.abs=final.com.abs%>%group_by(profid)%>%summarise(ncom = n(),rating = round(mean(rating, na.rm=T
          ),2),A1=round(mean(A.1, na.rm=T),2),A2=round(mean(A.2, na.rm=T),2),A3=round(mean(A.3, na.rm=T),2),
          A4=round(mean(A.4, na.rm=T),2),A5=round(mean(A.5, na.rm=T),2))
    rst=final.pf.abs
  }
  return(rst)
}
```