



Aspect-Based Sentiment Analysis for Sentence Types with Implicit Aspect and Explicit Opinion in Restaurant Review Using Grammatical Rules, Hybrid Approach, and SentiCircle

Suhariyanto¹ Rachmad Abdullah² Riyanarto Sarno^{1*} Chastine Fatichah¹

¹Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

²Department of Information Technology Management, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

* Corresponding author's Email: riyanarto@if.its.ac.id

Abstract: Sentiment analysis can provide rough recommendations in the form of sentiment from a collection of reviews or can provide recommendations in more detail about sentiment in a particular aspect called aspect-based sentiment analysis (ABSA). Sentiment analysis based on many aspects has been carried out but its accuracy is still being developed. In previous research, most research was carried out on explicit and implicit aspects and opinions and was carried out in simple sentences. The purpose of this research is to analyze the sentiment of restaurant reviews using the rule grammar method to extract implicit aspects - explicit opinions in four sentence models, namely simple (Si-AIOE), compound (Co-AIOE), complex (Ce-AIOE), and compound-complex (CoCe-AIOE). The ABSA method is proposed using the development of a grammatical rule extraction method to extract explicit and implicit aspect words and opinion words as the basis for sentence extraction. Rules making is done to take explicit and implicit aspect words and opinion words in Si-AIOE, Co-AIOE, Ce-AIOE, and CoCe-AIOE so that the comparison of the evaluation values can be known. This research uses the Semeval 2015 dataset on Restaurant reviews from the Tripadvisor Website which has been annotated as sentence data for ABSA. The aspect categorization process is then used to categorize aspects into 4 aspect categories, namely Ambience, Food, Service, and Price using hybrid approach. The hybrid approach is combined using Elmo-Wikipedia, grammatical rule extraction, WordNet, TF-ICF, and semantic similarity methods. The results of the aspect extraction obtained value of precision, recall, and f1-measure of 0.80, 0.84, and 0.82, respectively. Meanwhile, the ABSA process uses SentiCircle to classify sentiments into two, namely positive and negative. The results of the ABSA showed that the performance of proposed method achieve for precision, recall, and f1-measure were 0.84, 0.89, and 0.87, respectively.

Keywords: Aspect-based sentiment analysis, Grammatical rule extraction, Aspect categorization, Elmo-wikipedia, WordNet, TF-ICF, SentiCircle.

1. Introduction

The information age generates a lot of data in cyberspace, including product reviews from customers. One example of a review online is a TripAdvisor review. Meanwhile, the data format written in cyberspace is generally unstructured which must be processed and analyzed with the right technique [1]. One such technique is sentiment analysis. Sentiment analysis is a branch of science that analyzes public comments or opinions about a product, service, topic or organization with the aim of

to obtain useful information. Sentiment analysis is generally categorized into three categories: document level, sentence level, and aspect-based level [2]. Aspect-based sentiment analysis (ABSA) is an approach to find out a user's opinion from what aspects he likes and dislikes [3].

In previous researches, ABSA was applied in many explicit cases both aspects and sentiments. The research also focuses on cases of simple sentences and has not been able to solve well cases of other sentences such as compound, complex, and compound-complex. Hu and Liu [3] explain the simple sentence "The pictures are very clear" in

which the sentence is a case of an explicit sentence containing 1 explicit aspect and 1 explicit opinion.

In another research conducted by Reza [4] and Ayu [5] have worked on a similar case for explicit cases. These researches can take the hidden topic from the extracted aspect word so that it can expand the meaning of the aspect term to the aspect category in the review. For instance in one of their research can distinguish the aspect term "bass" for the 2 sentences "We can hear the bass sound" and "she like grilled bass". However, this research is only limited to works on explicit sentence cases, not for implicit cases. The latest research on the case of explicit sentences was conducted by Firmanto [6] who worked on the case of explicit sentences in compound sentences "the food is great and reasonably priced". The previous explicit sentence case research has not yet discussed sentence extraction based on other sentence structures.

The implicit aspect research has been done before by B Wang [7] who worked on the case of the sentence "It's beautiful and fashionable" using the linguistic rule method. The next aspect of implicit research was conducted by Z Hai [8], "too expensive, I really cannot afford it". Building on the previous case, this research worked on complex implicit sentences. However, the sentences that are done are only single adjective and single adverb as aspect term. W Zhang [9] in a later research worked on the implicit case of the aspect of "The price of this body wash is very expensive" in which there are two aspects of terms and one opinion. Z Hai [10] is working on a more in-depth implicit case, "The exterior is really attractive, and the screen is quite large, I like it very much!" where there are two aspects of the term and two adjectives as opinions. However, this research only works on the compound sentence pattern in the direction of the "and" conjunction, while the "but" conjunction has not been discussed. Related research about implicit aspect extraction has also been carried out by Tubishat [11]. This research carried out implicit case by combined implicit aspect corpus with dictionary based approach. However, all previous implicit sentence case research also did not yet discuss sentence extraction based on other sentence structures.

Bagus [12] in his research can lead to the development of an unsupervised method using machine learning to work on implicit cases such as "The land around here is flat" and "The party is bit flat". This research can distinguish the sentiment value of the word "flat" for each of these sentences. However, this research has not discussed the expansion of the meaning of opinion word that can allow differences in the value of sentiment polarity.

Based on previous research that has been done, this research carried out Aspect-Based Sentiment Analysis (ABSA) for restaurant reviews in cases of simple sentences that have implicit aspects and explicit opinions (Si-AIOE), cases of compound sentences with implicit aspects and explicit opinions (Co-AIOE), cases of complex sentences where there are implicit aspects and explicit opinions (Ce-AIOE), and cases of compound-complex sentences where there are implicit aspects and explicit opinions (CoCe-AIOE). The dataset used was obtained from the public dataset from Semeval 2015 [13] with a total of 2499 reviews. This dataset has been annotated, both the category aspect annotation and the sentiment polarity annotation.

This research raises the problem of extracting new implicit sentences that have not been done in previous researches. For example in the case of the following implicit simple sentence.

EX1: "the taste lasts long"

EX2: "the service was long"

"Long" on EX2 has a negative sentiment polarity while "long" on EX1 isn't always has a positive sentiment polarity. However, these problems have not been solved in previous researches. The word 'long' in EX1 and EX2 has different values of sentiment polarity because the context of the aspect of the sentence is different. EX1 contains the word "taste" which is a category of aspects of FOOD, while EX2 contains the word "service" which is a category of aspects of SERVICE.

Based on the implicit sentence extraction problems, both simple, compound, complex, and compound-complex, this research proposes an aspect categorization method using grammatical rule extraction and hybrid approach which consists of WordNet, TF-ICF, ELMo-Wikipedia, and Semantic similarity. This research aims to classify sentences into 4 sentence structures Si-AIOE, Co-AIOE, Ce-AIOE, and CoCe-AIOE. Furthermore, based on this structure, the content of the aspect word and opinion word in the sentence is taken.

Then the results of the aspect categorization process are used for the aspect-based sentiment analysis process. This aspect-based sentiment analysis research aims to differentiate sentiment polarity values on implicit aspects using the SentiCircle, Opinion Lexicon, Implicit Aspect Lexicon, and co-occurrence of context terms methods.

2. Related theory

Several theories related to the research will be explained in this section.

2.1 Aspect category keywords

This research used a restaurant review dataset from the TripAdvisor website. In the dataset, there are 4 (four) aspect categories reviewed, namely ambience, food, service, and price. These aspect categories are taken from the review of the aspect rating data on the TripAdvisor application. As has been explained about the definition of aspect categories [6] and the aspect category variables [14] so we used several keywords for four aspect categories in this research from Wikipedia. Table 1 below shows the aspect categories and variables that used in this research. Then Table 2 shows the aspect category keywords from Wikipedia that used in this research.

2.2 Pre-processing

In natural language processing, steps for pre-processing used are tokenizing → normalizing → punctuation removing → lemmatizing → stop words removing → correction of spelling.

Table 1. Aspect keyword variables

Aspect Categories	Variables
Ambience	<ul style="list-style-type: none"> • Atmosphere • Décor • furniture (table and doors) • environment (size, type, air conditioner) • entertainment (Tv, music, live performance) • places or areas (bathroom, smoking area, buffet, bar, dining room, rest room) <ul style="list-style-type: none"> • view, • located area • service staff appearance.
Food	<ul style="list-style-type: none"> • menu variety • food item (taste, quality, healthy, cooking level, size, religious dietary) <ul style="list-style-type: none"> • ingerdients • drink • dessert • appetizers.
Service	<ul style="list-style-type: none"> • Parking • all of staffs, managers and employees (behavior, experience, availability, appearance) <ul style="list-style-type: none"> • open hours <ul style="list-style-type: none"> • wifi • gift • delivery (time, ordering method) <ul style="list-style-type: none"> • seating.
Price	<ul style="list-style-type: none"> • Price value • reasonable price item.

Table 2. Aspect keyword from Wikipedia

Aspect Category	Wikipedia Page Titles
Ambience	Theme_restaurant Spirit_of_place
Food	Food Cuisine
Service	Customer_service Waiting_staff
Price	Price Value_(economics)

2.3 Grammatical rules

In this research, Grammatical Rule is used for sentence extraction to identify and split the sentences become Si-AIOE, Co-AIOE, Ce-AIOE, and CoCe-AIOE before implement the rule for extract aspect and opinion terms based on sentence structures. The Grammatical rule uses opinion lexicon and implicit opinion lexicon for taking the aspect word and opinion word which identified as Noun Phrase and Adjective Phrase.

2.4 Hybrid approach

In this research, we combine several methods to get best accuration score for aspect categorization process, there are Elmo, WordNet, TF-ICF, and Semantic similarity.

2.4.1. Elmo

Elmo is deep contextualized word representation [15]. The following is an example of explaining the implementation of the word "bucket" below:

E1: "The new lamp had good light for reading."

E2: "He likes to travel light."

E3: "Magnesium is a light metal."

The word "light" from each sentences of E1, E2, and E3 has different meaning to each other. Based on three example sentences, Elmo can generate different vectors for the word "light" in each sentence depends on sentence context. This research uses TensorFlow tools [16] for Elmo implementation process.

2.4.2. WordNet

In this research, WordNet [17] is used to expand aspect and opinion words that extracted from their synonyms and antonyms.

2.4.3. TF-ICF

TF-ICF [18] is used in this research to expand the aspect and opinion words that extracted from calculating the word frequency from the document on

each cluster. Eq. (1) shows the ICF equation. Then Eq. (2) shows the TF-ICF equation.

$$ICF_x = \log\left(\frac{C}{CF_x}\right) \quad (1)$$

$$TF - ICF_{x,i} = TF_{x,i} \times ICF_x \quad (2)$$

Where:

- x = Word term
- i = Cluster
- C = Cluster total
- CF_x = Cluster frequency total of word term x
- $TF_{x,i}$ = Frequency total of word term x in cluster i

2.4.4. Semantic similarity

Semantic similarity [19] in this research is using cosine similarity [20] to calculate two words similarity based on meaning. Eq. (3) shows the equation of cosine similarity.

$$cosine(w_a, w_b) = \frac{\sum_{i=1}^n w_{ai} w_{bi}}{\sqrt{\sum_{i=1}^n (w_{ai})^2} \sqrt{\sum_{i=1}^n (w_{bi})^2}} \quad (3)$$

where,

- w_a = word 1
- w_b = word 2
- w_{ai} = vector member from w_a
- w_{bi} = vector member from w_b

2.5 Opinion lexicon

In this research, lexicon for explicit sentence case is used from Qiu [21] that developed Opinion Lexicon from Hu [3] became around 6800 words. This lexicon contains positive opinion lexicon and negative opinion lexicon.

2.6 Implicit aspect lexicon

This research uses the implicit aspect corpus [22] to extract aspect categories with implicit sentence cases.

2.7 SentiCircle

SentiCircle [23], in line with the distributional hypothesis [24] about similar words, in this research is used to captures the words based on Semantics context and concept. Then, sentiment word polarity of word context is calculated using SentiMedian. The SentiMedian equation as follows in Eq. (8). The result is groping into 4 kwadran, namely: positive,

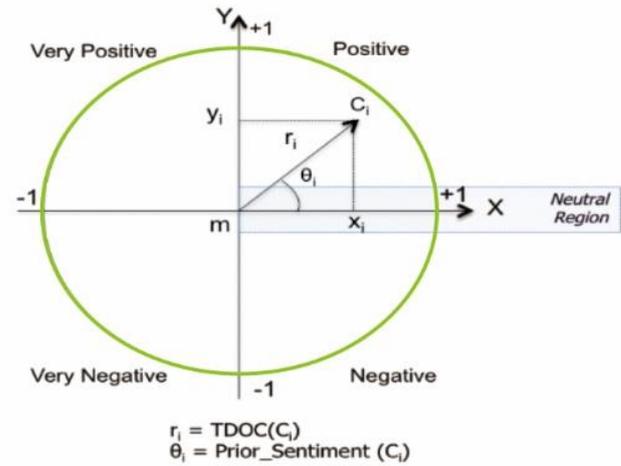


Figure. 1 SentiCircle model

very positive, negative, or very negative. SentiCircle system model is illustrated in Fig. (1). This model, with polar coordinate system, assumes the process as follows.

1. Generate the correlation term degree ($TDOC$)
2. Calculate polarity value with sentiment lexicon.
3. Calculate point x_i and y_i .

$$TDOC(m, c_i)d = f(c_i, m) \times \log \frac{T}{T_{c_i}} \quad (4)$$

$$\theta_i = Prior_Sentiment(c_i) * \pi \quad (5)$$

$$x_i = r_i \cos \theta_i \quad (6)$$

$$y_i = r_i \sin \theta_i \quad (7)$$

$$g = \min_{g \in R^2} \sum_{i=1}^n \|p_i - g\|^2 \quad (8)$$

where,

$TDOC(m, c_i)d$ = term degree of correlation between word m and word context c_i in document d

T = total m in d

T_{c_i} = total c_i in d

$f(c_i, m)$ = frequency of joint occurrences between m and c_i in d

$Prior_Sentimen$ = polarity value

θ_i = c_i degree (in radian)

x_i = c_i position in x axis

y_i = c_i position in y axis

p_i = c_i position

g = position of m (x_m and y_m)

2.8 Evaluation

Testing aspect categorization and sentiment analysis uses three evaluation methods, namely

Table 3. Confusion matrix

		Actual Values	
		Positive	Negative
Predicted Values	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Precision, recall, and F1-measure. Table 3 shows confusion matrix to calculate this evaluation.

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

$$Recall = \frac{TN}{TP+FN} \tag{10}$$

$$F1 - measure = \frac{2 * P * R}{P + R} \tag{11}$$

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{12}$$

3. Research method

This research begins with pre-processing method to take data without noise or less relevant information. The pre-processing result is used as the input in aspect categorization module. We input aspect keywords from Wikipedia links using Elmo. We use

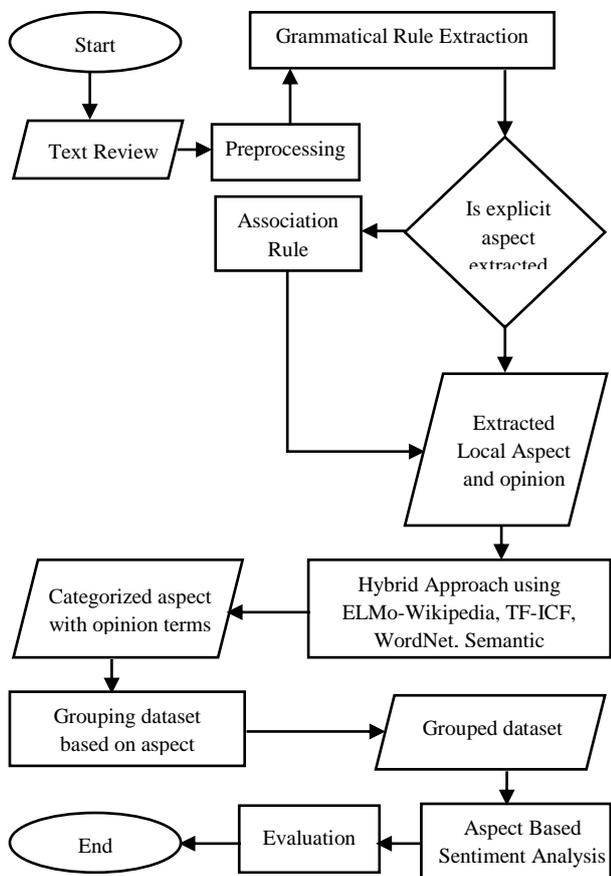


Figure. 2 System flow diagram

it to take aspect and opinion word terms using grammatical rule extraction. Then the aspect categorization process to extract the aspect and opinion words using hybrid approach. Hybrid approach is combined using comparison between Elmo-Wikipedia result, grammatical rule extraction result, WordNet, TF-ICF, and Semantic similarity. The result of hybrid approach process is categorized aspect with opinion terms.

Thus we prepare the grouped dataset for aspect-based sentiment analysis process using SentiCircle. Finally, grammatical rule extraction, aspect categorization and aspect-based sentiment analysis are measured and evaluated by evaluation method proposed in this research.

3.1 Dataset

The dataset is obtained from TripAdvisor website [13]. Total dataset in this research shows in Table 4. We split the reviews based on structure sentence into several sentences. Then several sentence is labelled based on ID Review and ID Sentence to different review from customers. The review data for dataset representation shows in Table 5.

3.2 Pre-processing

The review pre-processing results are words, lower case, removing punctuation, removing stop words, reducing, and the minimum word limit.

3.3 Sentence extraction using grammatical rule extraction

Review that has been done preprocessing will take a pair of aspect words and opinions using grammatical rule extraction. In this process, the basic dependencies parser and the Enhanced++

Table 4. Dataset representation

Domain	Train	Test	Total
Restaurant	1654	845	2499

Table 5. Review data for dataset representation

ID Review	ID Sentence	Review
1004293	1	We, there were four of us, arrived at noon - the place was empty - and the staff acted like we were imposing on them and they were very rude.
1004293	2	They never brought us complimentary noodles, ignored repeated requests for sugar, and threw our dishes on the table.

dependencies parser from Stanford CoreNLP are used. In the initial phase of sentence extraction, the expert succeeded in to ago classify 2499 sentences into 4 types of sentence structure. The detailed steps for each process are as follows:

1. Splitting sentence

In this module, the process of solving simple, compound, complex sentences, and compound-complex sentences will be carried out.

2. Aspect and opinion word extraction

After breaking it down into 4 sentence structures, the next process is to take aspect words and opinion words Table 7, Table 8, Table 9, and Table 10 shows the results of testing the data processing by the grammatical rule extraction system for Si-AIOE, Co-AIOE, Ce-AIOE, CoCe-AIOE respectively.

This research uses grammatical rule extraction that develops the rule from Poria et al [11]. It used to extract aspect and opinion words. There are two basic general rules for proposed aspect parser:

- Rules for sentences having a Noun Phrase (NP)
- Rules for sentences do not have a Noun Phrase (NP) and have an Adjective Phrase (AP)

Table 6. Result of Si-AIOE extraction

ID Review	ID Sentence	Review
1205520	3	It's delicious!
FF#6	0	So rude!!!
WE#1	0	Expensive
...

Table 7. Result of Co-AIOE extraction

ID Review	ID Sentence	Review
1145510	2	But the pizza is way too expensive
ADLT#6	5	His response was smug, arrogant, and condescending, totally consistent with his department on display all evening.
BFC#4	0	Overpriced and not tasty
...

Table 8. Result of Ce-AIOE extraction

ID Review	ID Sentence	Review
512294	1	If you are looking for good quality, cheap eats - this is the place.
ZCeIE#1188	0	If you're impulsive like me, then \$6 is ok
ZCeIE1192	0	I loved that these were so delicious
...

Table 9. Result of CoCe-AIOE extraction

ID Review	ID Sentence	Review
ZCoCeIE#1101	0	The first one that came to me was so delicious and feeling grateful
ZCoCeIE#1102	0	This restaurant was full and rather noisy since we came
ZCoCeIE#1103	0	There was uncomfortable, however, they were incredibly unwelcoming and arrogant.
...

3.4 Expanding keyword terms

The expanding algorithm of keyword terms extraction using WordNet and TF-ICF can be seen in Fig. 3.

3.5 Aspect categorization (AC)

In the aspect categorization stage, the aspect category is calculated using semantic similarity.

1. Retrieve keywords data from Wikipedia
2. Pre-process data as an input
3. Retrieve synonym and antonym of word terms from WordNet
4. Calculate value of ICF with Eq. (1)
5. Calculate value of TF-ICF with Eq. (2)
6. Compare ten highest similarity value of word terms with lexicons
7. Extract the highest value of word term as expanded keyword terms.

Figure. 3 Expanding process for keyword terms

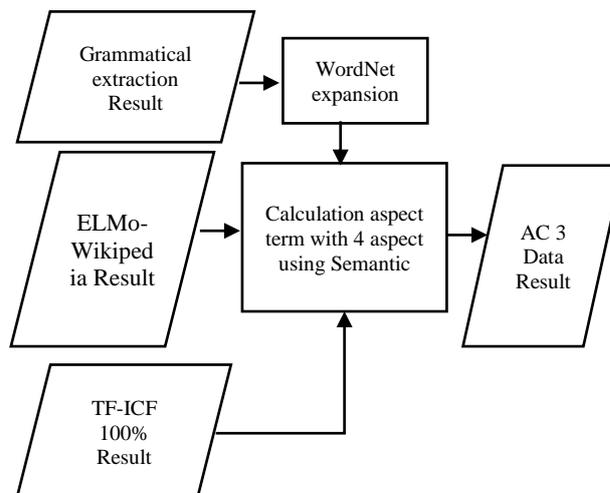


Figure. 4 AC3 process

Semantic similarity calculates the aspect keywords which determined using Elmo-Wikipedia and aspect terms which determined using TF-ICF. The aspect categorization process is divided into 3 stages, namely AC1, AC2, and AC3. AC1 method using Grammatical Rule, ELMo-Wikipedia, and Semantic. AC2 method using Grammatical Rule, ELMo-Wikipedia, TF-ICF 20%, and Semantic. AC3 method using Grammatical Rule, WordNet, ELMo-Wikipedia, TF-ICF 100%, and Semantic.

3.6 Aspect based sentiment analysis (ABSA)

In this stage, we conduct three experiments to determine the best Aspect-Based Sentiment Analysis (ABSA) performance. SentiCircle method is used on ABSA1. ABSA2 using SentiCircle and Opinion Lexicon. ABSA3 using SentiCircle, Opinion Lexicon, Implicit Aspect Lexicon, and Co-occurrence of context terms.

4. Result and analysis

4.1 Result of grammatical rule extraction

The evaluation results of grammatical rule extraction in every structure show in Table 10. Table 10 shows that an accurate score of grammatical rule extraction can extract well for aspect words and opinion words in 4 types of sentence structures, For instance, one of review in Si-AIOE is “it’s delicious!”; the aspect word is implicitly shown to the food has been predicted by an expert. Grammatical rule extraction detected the sentence does not has POS an NP and has an AP so rule-based algorithm work to predicted aspect word using implicit aspect lexicon. The result is detected delicious labelling as quality from Food Category.

4.2 Result of AC

This stage shows that WordNet and TF-ICF 100% can expand the existing keyword terms from Wikipedia. They can predict aspect and opinion word more precisely. For instance as shown in Table 11, review “unhygienic”; this review could not be predicted by AC1 because Wikipedia only finds the basic word 'hygiene'. AC3 method can predict correctly using WordNet to expands the synonym and antonym of ‘hygiene’. The expand term list of ‘unhygiene’ which exist in opinion lexicon is ‘unhealthy’. Then TF-ICF expands with calculating the term list in the document. So, ‘unhealthy’ result is labelled as a variable of food category in aspect categorization process. For more detail, the evaluation score of aspect categorization

Table 10. Accuration of grammatical rule extraction

Structure	Accuracy Score	
	Aspect word	Opinion word
Si-AIOE	0.87	0.90
Co-AIOE	0.77	0.88
Ce-AIOE	0.81	0.86
CoCe-AIOE	0.71	0.81

Table 11. AC process comparison for word “unhygienic”

AC1 Keyword	Semantic similarity	AC result
Null	Null	Null
AC3 Keyword	Semantic similarity	AC result
Unhealthy	-0.5	Food

Table 12. Result of AC performances

AC Performances		
AC Approach	Method	F1-measure
AC1	Grammatical Rules + ELMo-Wikipedia + Semantic	0.71
AC2	Grammatical Rules + WordNet + ELMo-Wikipedia + TF-ICF 20% + Semantic	0.76
AC3	Grammatical Rules + WordNet + ELMo-Wikipedia + TF-ICF 100% + Semantic	0.82

performances are showed in Table 12. Then, result of every performance of aspect categorization shows in Table 13. Table 13 shows the results of sentences extraction by comparing the result of the system than expert for Si-AIOE, Co-AIOE, Ce-AIOE, CoCe-AIOE respectively. Score 0 if the word which doing by an expert is 1 and the predict which doing by an algorithm is not 1, else the score is 1.

4.3 Result of ABSA

In this stage, the evaluation of performance results shows in Table 14. Table 14 shows a comparison of 3 performance using SentiCircle, there are ABSA1, ABSA2, and ABSA3. SentiCircle is used in ABSA1 and ABSA2 to predict the context of the word in explicit case very well. For instance, these are result of 4 review to distinguish the sentiment of a word and its polarity.

1. The wait here is long for dim sum, but if you don't like sharing tables or if the typical raucous dim sum atmosphere is not your gig, this is a sleek (for Chinatown) alternative.
2. Service was long, made me disappointed.

Table 13. Sentence categories extraction

ID Review	ID Sent	Extraction Result (doing by algorithm)	Polarity (doing by algorithm)	Aspect			Opinion		
				word (doing by expert)	predict (doing by algorithm)	score true = 1; false = 0	word (doing by expert)	predict (doing by algorithm)	score true = 1; false = 0
Si-AIOE									
1205520	3	delicious_delicious	positive	-	delicious	1	delicious	delicious	1
FF#6	0	rude !rude	negative	-	rude	1	rude	rude	1
WE#1	0	expensive !expensive	negative	-	expensive	1	Expensive	expensive	1
Co-AIOE									
1145510	2	pizza!expensive	negative	pizza	pizza	1	expensive	expensive	1
ADLT#6	5	response!smug response!arrogant response!condescending	negative	Response	response	1	Smug, arrogant	smug, arrogant	1
BFC#4	0	Overpriced !overpriced tasty !tasty	negative	-	overpriced, tasty	1	Overpriced, tasty	overpriced, tasty	1
Ce-AIOE									
512294	1	Eats !good eats!cheap	positive	Eats	eats	1	Good, cheap	good, cheap	1
ZCeIE#188	0	\$!ok	positive	\$6	\$6	1	Ok	ok	1
ZCeIE#1192	0	Delicious! delicious	positive	-	delicious	1	Delicious	delicious	1
CoCe-AIOE									
ZCoCeIE#1101	0	Delicious! delicious	positive	-	delicious	1	Delicious	delicious	1
ZCoCeIE#1102	0	Restaurant! Full restaurant !noisy	negative	Restaurant	restaurant	1	Full, noisy	full, noisy	1
ZCoCeIE#1103	0	Unwelcoming !unwelcoming arrogant !arrogant.	negative	-	unwelcoming, arrogant	1	Unwelcoming, arrogant	unwelcoming, arrogant	1

Table 14. Result of ABSA performances

ABSA Performances		
ABSA Approach	Method	F1-Measure
ABSA1	SentiCircle	0.73
ABSA2	SentiCircle + Opinion Lexicon	0.79
ABSA3	SentiCircle + Opinion Lexicon + Implicit Aspect Lexicon + Cooccurrence of context terms	0.87

3. The taste lasts long and I'm very satisfied.
 4. feels very long on the tongue, very delicious.
 SentiCircle can predict context of the word in explicit case very well. For instance, 'long' in review 1 and review 2 as negative sentiment from aspect

word 'service' because 'long' is detected as negative

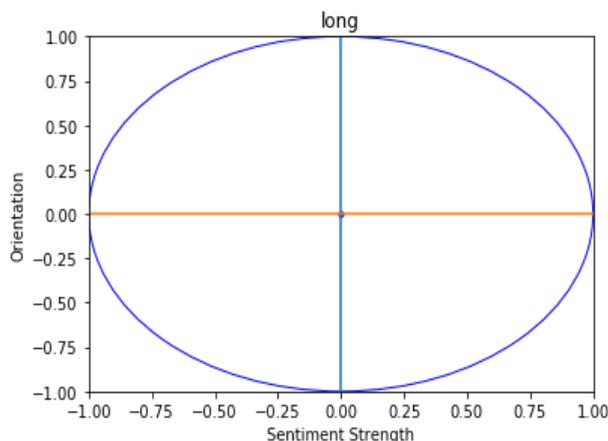


Figure. 5 Ambience

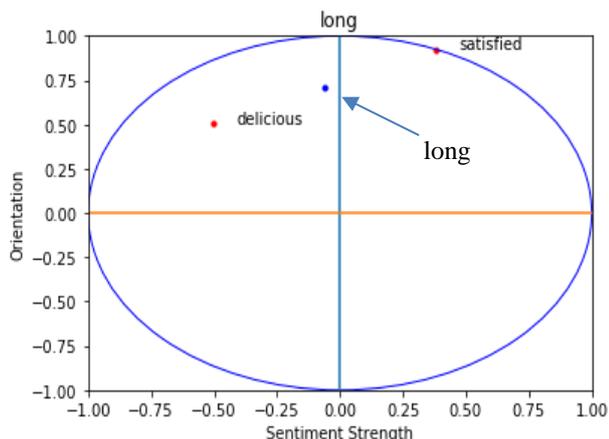


Figure. 6 Food

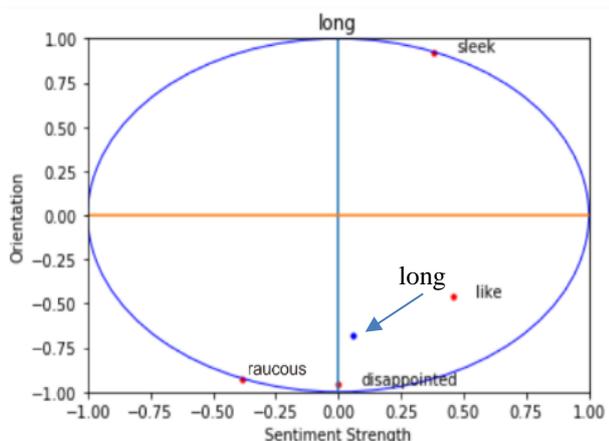


Figure. 7 Service

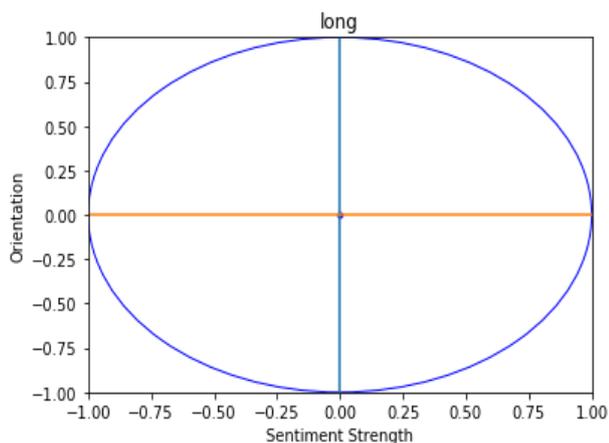


Figure. 8 Price

lexicon in Opinion Lexicon. While the context of the word is the implicit case, it does not give the best performance like explicit case. For instance, review 3 and review 4, 'long' could predict as negative sentiment using ABSA1 and ABSA2. ABSA3 give different value because 'long' will extract as adjective modifier from implicit aspect lexicon 'lasts' that in co-occurrence with 'taste'. Thus, ABSA3 predict that's review as a positive sentiment.

Table 15. Comparison of ABSA

Review	Opinion Extracted	ABSA result	
		AC	Sentiment
It's delicious!	delicious	ABSA1	
		Null	Null
		ABSA2	
		Null	Null
		ABSA3	
		Food	Positive
But the pizza is way too expensive.	expensive	ABSA1	
		Food	Positive
		ABSA2	
		Food	Negative
		ABSA3	
		Price	Negative
If you are looking for good quality, cheap eats - this is the place.	good cheap	ABSA1	
		Food	Negative
		ABSA2	
		Food	Positive
		ABSA3	
		Price	Positive
The first one that came to me was so delicious and feeling grateful.	delicious grateful	ABSA1	
		Null	Null
		ABSA2	
		Null	Null
		ABSA3	
		Food	Positive

5. Conclusion

This research proposes aspect based sentiment analysis for sentence types with implicit aspect and explicit opinion using grammatical rules, hybrid approach, and SentiCircle. Grammatical rules extraction is used to extract and split the sentences to get aspect and opinion word. Hybrid approach method is used to calculate word similarity in aspect categorization module. Then SentiCircle, opinion lexicon, implicit aspect lexicon, and co-occurrence of context term method are used to aspect based sentiment analysis.

AC3 performance value with F1-Measure of 0.82 show that AC3 is the best performance than AC1 and AC2. The best performance for aspect-based sentiment analysis is ABSA3 with F1-Measure of 0.87. ABSA3 expands word context of SentiCircle using opinion lexicon, implicit aspect lexicon, and co-occurrence word.

SentiCircle change the review sentiment polarity based on context in explicit case properly but not in implicit case. While in implicit case, SentiCircle which combined with opinion lexicon, implicit aspect lexicon, and co-occurrence of context terms could work properly. Experiment polarity of word sentiment "long" for FOOD category which different

with SERVICE category show how to SentiCircle solve two problems of sentence type cases.

For future work, the method for aspect categorization and sentiment analysis can be enhanced to solve sentence that contains an explicit aspect and implicit opinion and sentence that contains an implicit aspect and implicit opinion.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

This research can work well and successfully because of the following research contributions: Conceptualization by Suharyanto and Prof. Riyanarto Sarno; methodology and software by Suharyanto and Rachmad Abdullah; validation, formal analysis, investigation, resources by Prof. Riyanarto Sarno, Suharyanto, and Rachmad Abdullah; data curation, writing-original draft preparation by Suharyanto; writing-review and editing, visualization, and supervision by Prof Riyanarto Sarno and Chastine Fatichah; project administration and funding acquisition by Suharyanto.

Acknowledgments

This research was funded by the Indonesian Ministry of Education and Culture under Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) Program managed by Institut Teknologi Sepuluh Nopember (ITS) and under Riset Inovatif-Produktif (RISPRO) Invitation Program managed by Lembaga Pengelola Dana Pendidikan (LPDP).

References

- [1] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics", *Int. J. Inf. Manage.*, 2015.
- [2] B. Liu, "Sentiment Analysis and Opinion Mining", *Synth. Lect. Hum. Lang. Technol.*, Vol. 5, No. 1, pp. 1-167, 2012.
- [3] M. Hu and B. Liu, "Mining and summarizing customer reviews", In: *Proc. of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '04*, p. 168 2004.
- [4] R. Priyantina and R. Sarno, "Sentiment Analysis of Hotel Reviews Using Latent Dirichlet Allocation, Semantic Similarity and LSTM", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 4, pp. 142-155, 2019.
- [5] D. Ayu and K. Khotimah, "Sentiment Analysis of Hotel Aspect Using Probabilistic Latent Semantic Analysis, Word Embedding and LSTM", Vol. 12, No. 4, 2019.
- [6] A. Firmanto and R. Sarno, "Aspect-based sentiment analysis using grammatical rules, word similarity and SentiCircle", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 5, pp. 190-201, 2019.
- [7] B. Wang and H. Wang, "Bootstrapping Both Product Features and Opinion Words from Chinese Customer Reviews with Cross-Inducing 1", *Science (80-.)*, No. 60675035, pp. 289-295, 1998.
- [8] Z. Hai, K. Chang, and J. J. Kim, "Implicit feature identification via co-occurrence association rule mining", *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Vol. 6608 LNCS, No. PART 1, pp. 393-404, 2011.
- [9] W. Zhang, H. Xu, and W. Wan, "Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis", *Expert Syst. Appl.*, Vol. 39, No. 11, pp. 10283-10291, 2012.
- [10] Z. Hai, K. Chang, G. Cong, and C. C. Yang, "An association-based unified framework for mining features and opinion words", *ACM Trans. Intell. Syst. Technol.*, Vol. 6, No. 2, 2015.
- [11] M. Tubishat and N. Idris, "Explicit and Implicit Aspect Extraction using Whale Optimization Algorithm and Hybrid Approach", In: *Proc. of International Conference on Industrial Enterprise and System*, pp. 208–213, 2019, doi: 10.2991/icoiese-18.2019.37.
- [12] B. S. Rintyarna, R. Sarno, and C. Fatichah, "Semantic Features for Optimizing Supervised Approach of Sentiment Analysis on Product Reviews", *Computers*, vol. 8, no. 55, pp. 1–16, 2019, doi: 10.3390/computers8030055.
- [13] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, "SemEval-2015 Task 12: Aspect Based Sentiment Analysis", In: *Proc. of International Workshop on Semantic Evaluation*, pp. 486–495, 2015, doi: 10.18653/v1/s15-2082.
- [14] F. Nurifan, R. Sarno, and K. R. Sungkono, "Aspect Based Sentiment Analysis for Restaurant Reviews Using Hybrid ELMO-Wikipedia and Hybrid Expanded Opinion Lexicon-SentiCircle", *Int. J. Intell. Eng. Syst.*, Vol. 12, No. 6, pp. 47-58, 2019.
- [15] M. E. Peters, M. Neumann, M. Lyyer, M. Gardner, C. C;ark, K. Lee, and L. Zettlemoyer, "Deep Contextualized Word Representations", In: *Proc. of North American Association for Computational Linguistics*, pp. 2227–2237, 2018, doi: 10.18653/v1/n18-1202.

- [16] M. Abadi, "TensorFlow: A system for large-scale machine learning", In: *proc. of 12th USENIX Symp. Oper. Syst. Des. Implementation, OSDI 2016*, pp. 265-283, 2016.
- [17] C. Fellbaum, "WordNet(s)", in *Encyclopedia of Language & Linguistics*, Elsevier, 2006, pp. 665-670.
- [18] L. H. Suadaa and A. Purwarianti, "Combination of Latent Dirichlet Allocation (LDA) and Term Frequency-Inverse Cluster Frequency (TFxICF) in Indonesian text clustering with labeling", In: *Proc. of 2016 4th International Conference on Information and Communication Technology, ICoICT 2016*, pp. 1-6, 2016.
- [19] J. M. Gil, "An overview of textual semantic similarity measures based on web intelligence", *Artif. Intell. Rev.*, Vol. 42, No. 4, pp. 935-943, 2014.
- [20] P. Xia, L. Zhang, and F. Li, "Learning similarity with cosine similarity ensemble", *Inf. Sci. (Ny)*, Vol. 307, pp. 39-52, 2015.
- [21] G. Qiu, B. Liu, J. Bu, and C. Chen, "Expanding domain sentiment lexicon through double propagation", *IJCAI Int. Jt. Conf. Artif. Intell.*, No. November 2014, pp. 1199-1204, 2009.
- [22] I. Cruz, A. Gelbukh, and G. Sidorov, "Implicit Aspect Indicator Extraction for Aspect-based Opinion Mining", *Int. J. Comput. Linguist. Appl.*, Vol. 5, No. 2, pp. 135-152, 2014.
- [23] H. Saif, M. Fernandez, Y. He, and H. Alani, "SentiCircles for contextual and conceptual semantic sentiment analysis of Twitter", in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014.
- [24] P. D. Turney and P. Pantel, "From frequency to meaning: Vector space models of semantics", *J. Artif. Intell. Res.*, Vol. 37, pp. 141-188, 2010.