International Journal on Advanced Science Engineering Information Technology

Vol.12 (2022) No. 4
ISSN: 2088-5334

Aspect Based Opinion Summarization Using Rule-Based Method and Support Vector Machine for Indonesian Reviews

Pindya Khabiibah [a,*], Sari Widya Sihwi [a], Haryono Setiadi [a]

[a] *Department of Informatics, Universitas Sebelas Maret, Surakarta, Indonesia*
Corresponding author: [*]pindyakhabiibah@student.uns.ac.id

***Abstract*—E-commerce's reviews feature can help users find information about the desired skin care products to choose the right one. However, the reviews number of a product grows rapidly due to the popularity of e-commerce and the product itself. A user becomes difficult to read all reviews one by one and extract useful information. To deal with this problem, we summarize aspects using the Rule-Based method and Support Vector Machine. We propose a Rule-Based method that is used to break down a review into several segments based on its aspect. Support Vector Machine is used to classify sentence segments according to their polarity. The data used in this study is Indonesian reviews of skin care products obtained from the Female Daily website. The average accuracy results using 10-fold cross-validation of sentiment classification is 74%. We experimented on 462 reviews where the accuracy is 92% in aspect categorization and 71.2% in sentiment classification. Based on humans, the lowest value is the suitability of a sentence with its sentiment/polarity. The highest value is the suitability of sentences with its aspect and usability of summary to helps users to find specific information so they can decide whether to buy that product or not. It can be concluded that the reader can well receive the summary. Future work can consider the negation word to reduce misclassification in the sentiment classification step.**

***Keywords*— Opinion summarization; product review; rule-based; support vector machine.**

Manuscript received 28 Jul. 2020; revised 15 Mar. 2021; accepted 23 Apr. 2021. Date of publication 31 Aug. 2022.
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

I. INTRODUCTION

Nowadays, it is common for users to read others' opinions (reviews) to find information about a product before buying it in an e-commerce [1]. It becomes mandatory for a customer since all producers or sellers offer many claims on their products, such as a skin care product will be promoted based on its effect, substance, variant, and specific technology [2]. However, it is not an easy task to read and analyze others' reviews on the targeted product since the huge amount of the review [1], [3], noisy text [3], [4], redundancy [4] and irrelevant content of the review [1], [5]. Thus, opinion summarization of the existing reviews is required to get the important information of the product's aspects, including their descriptions and polarities [6].

One of the main approaches to generate opinion summaries is aspect-based opinion summarization. This approach generates summaries of opinions for the main evaluated aspects of an entity. The idea comes from the true product reviews for a user: its aspects and sentiments [6]. Therefore, aspect-based opinion summarization is quite different from traditional textual summary [1], which implies making an opinion summary.

The aspect-based opinion summarization is one of the main approaches to generating opinion summaries with three general steps [4], [7]. The first step is aspect identification that can be done by brainstorming and interviews, extracting from reviews, clustering, or taking from the literature [8]. The second step is sentiment prediction or sentiment analysis with two approaches, Lexicon Analysis or Machine Learning [9]. The third step is generating a concise and digestible summary of opinions, in the forms of structural, visual, or textual [1], [3], and use an extractive approach or abstractive approach [10]. An extractive summary is generated by selecting the sentence or sentence segment that best represents the original opinion, and abstractive summarization is generated by rewriting using new sentences as summary content [1].

Most studies in opinion summarization using aspect-based used English text in various domains, such as tourism [5], [11], mobile phones [12], and many more and also with various methods. Since aspect-based opinion summarization has three general steps, various methods can be used for each step. A study by Yauris and Khodra [13] proposed a modified Double

Propagation. Condori and Pardo [1] compared an extractive method, Opizer-E, with an abstractive method, Opizer-A, which both uses lexicon analysis to generate the polarities. Ramadhan *et al.* [12] used TextRank, a graph-based extractive text summarization algorithm, and did not generate the polarity of each review. Tran *et al.* [3] proposed Automatic Aspect-based Sentiment Summarization (AAbSS) and manifested the results in structural, visual, and textual.

We do not find publication on aspect-based opinion summarization using Indonesian text based on our search on the Internet. Most research stops their works until performing sentiment analysis, without presenting a summary of overall opinions and without considering aspects of the entity [14]–[18]. Few others are only taking Most researches in Indonesian domain, only take one or two-part of general steps of aspect-based opinion summarization as their purpose, which cannot be said as aspect-based opinion summarization. For example, study Darmawiguna et al. [19] performs opinion summarization using the TextRank algorithm to summarize the reviews in the structural form. However, Darmawiguna *et al.* [19] have not taken into account aspects of the product in summarizing the opinions.

Another example is Ekawati and Khodra [20], which has considered aspects based on sentiment analysis but did not summarize. It used a rule-based algorithm to divide a sentence into some segments according to the aspect contained by each segment and also by conjunction or punctuation. However, because reviews were written in an informal language and format structure, we will still leave a segment with multi aspects if Ekawati and Khodra [20]'s rule-based algorithm is implemented.

This paper proposes a method to make an extractive opinions summary in the structural form by considering the aspects. Because each segment of a sentence will have aspects and a polarity (sentiment), we make sure each segment does not have multi-aspect and ambiguous polarity. This research uses Indonesian skin care product reviews taken from the website of Female Daily, e-commerce, editorial, and online community of Indonesian beauty enthusiasts. Because research on the topic using Indonesian language texts is still rarely done, we hope this research can contribute to the development of aspect-based review summarization research in Indonesian.

Following are the proposed approaches in this study. To identify aspects, we extract manually from all corpus filtered by the chosen minimum frequency. To divide a sentence into some segments, we propose a new Rule Based algorithm, which can separate a multi-aspects sentence becomes several segments, even though there is no conjunction or punctuation. We will implement sentiment prediction in segment level using Support Vector Machine (SVM) to have information on reviews polarity. SVM is chosen because it shows a good performance in doing classification [15], including in the cases of Indonesian [16]–[18], compared to other methods. To summarize the review, we used a sentence to sentence similarity feature to get consumers' comments and use a threshold to filter sentences based on similarity. Then, the length of the sentence feature is calculated to obtain the sentence with the sentence most information. Then, one sentence with the highest score will be chosen for the summary.

II. Material and Methods

The general steps on aspect-based summarization is expanded become six steps in our proposed methods. Those are Data Collection, Text Preprocessing, Aspect Based Sentences Segmentation, Using Rule-Based, Sentiment Classification Using Support Vector Machine, Sentences Selection & Summary Generation, and Summary Evaluation. The overall process can be depicted from Figure 1 and described as follow.

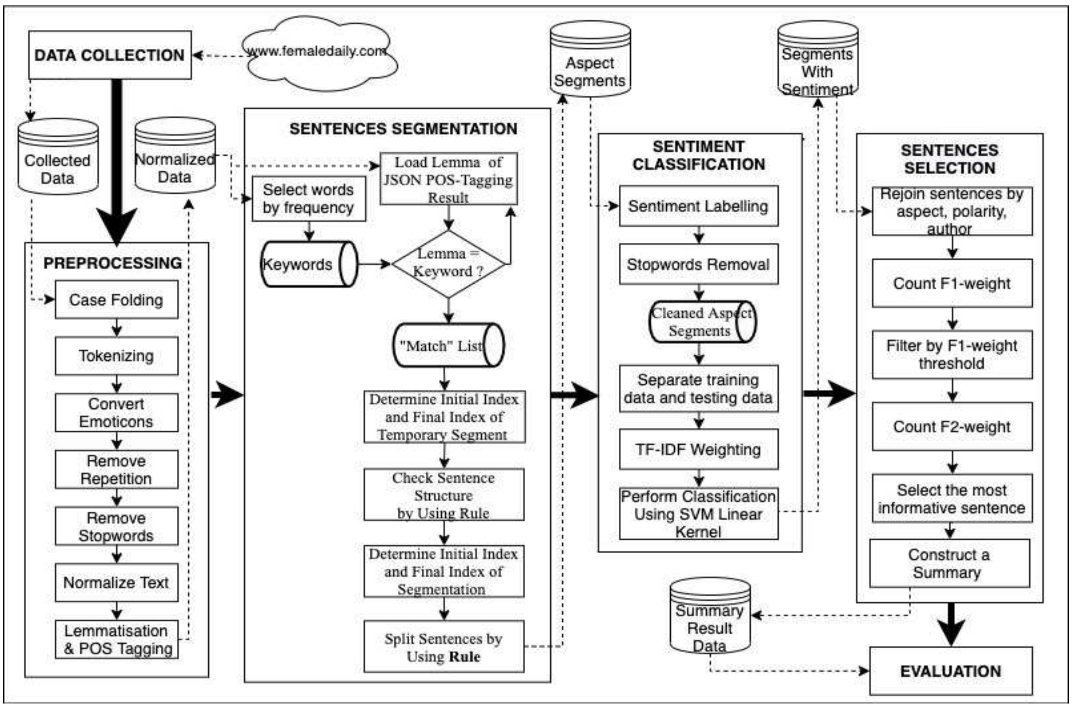

Fig. 1 Overall Process of Indonesian Aspect Based Opinion Summarization

A. Data Collection

At this stage, skin care product review data was taken from the Female Daily website (http://reviews.femaledaily.com) by crawling using python with the help of the Selenium Web Driver library and BeautifulSoap4. The data is a review of skin-soothing treatment category by selecting three popular products or products with the highest number of reviews. These products are Nature Republic Aloe Vera 92% Soothing Gel (Product-1), the Saem 99% Jeju Fresh Aloe Soothing Gel (Product-2), and Innisfree Aloe Revital Soothing Gel (Product-3). The review number of those products respectively are 2918, 826, and 462. The data taken is review data published from 2019 specifically for Product-1. As for Product-2 and Product-3, use data from 2017 to 2019. This is because the reviews of the two products are few.

B. Text Preprocessing

After the data is obtained, the next step is the preprocessing process to eliminate noise. In this research, text preprocessing takes several steps, including:

- *Case folding*, lowering text, and delete a non-ASCII character.
- *Tokenizing*, to separate text into a collection of tokens (words) *from* a sentence.
- *Convert emoticons*, convert emoticons according to their meaning, such as turn :) and <3 into smile and love.
- *Remove repetition* to delete characters that are written repeatedly. The example is from "*sayaaaaa pergi ke pasar" to "saya pergi ke pasar*".
- *Remove stopwords,* delete words that are not standard and interjection words (exclamation words) that are considered unimportant in a sentence.
- *Normalize text* to change non-standard words utilizing the list of word normalization that has been made [21], namely Colloquial Indonesian Lexicon and 968 words added by the author.
- *Lemmatization and POS tagging*, extract the base words and label them according to word classes using the API syntax analyzer from prosa.ai.

C. Aspect Based Sentences Segmentation Using Rule-Based

This stage aims to break down and group the sentences in the review that discuss the same aspects. In sentence segmentation, it consists of two main stages. There are the determination of aspects and keywords and separation of sentences based on aspects. The following is an explanation of the stages that will be carried out:

1) Aspect and Keywords Selection: In this study, nouns ('NNO' in tagset), verbs ('VBI', 'VBT', and 'VBP' in tagset), and adjectives ('ADJ' in tagset) were extracted from 4,352 Indonesian-language of skin care product reviews. Aspects and keywords are determined by calculating the frequency of each word in the product review. We sort them in descending order. Then, the selection of aspects and keywords is made manually. The list of frequencies selected as keyword candidates is a word that appears at least 15 times in all reviews.

2) Sentences Segmentation Based Aspects: At the stage of sentence segmentation, the Rule-Based method aims to break the sentence into several segments according to sentence patterns based on the corresponding aspects. Sentence segmentation is done by matching keywords with tokens in sentences, determining initial and final indices, checking sentence structure, and combining word tokens. Keywords that match with tokens are called windows.

Sentence's segmentation is done by considering the Part of Speech Tag (POS-Tag) of each token and the distance of the window with the token that supports it. Sentence's segmentation consists of three steps. There are matching keywords with sentences token, sentence's structure checking, and sentences segmentation.

Matching keywords with tokens in sentences are done to find out whether the tokens are the same as keywords. If the tokens are the same as the keywords, the tokens and aspect indexes are added to the list for later checking. Indexes and aspects stored in the list are data indexes, sentence indexes, keyword indexes, index of aspect groups, length of sentences, and appropriate keywords. This process' pseudocode is shown in Figure 2.

```
input = postag in json
if (lemma = noun keyword AND postag = 'NNO') OR
(lemma = verb keyword AND postag = 'VB') OR
(lemma = adjective keyword AND postag = 'ADJ'):
   add data index, sentence index, token index, aspect
   index, sentence length into window
```

Fig. 2 Keyword Matching

After the window matrix is obtained, a matrix row element is removed if sequential tokens contain the same aspects and they are from the same sentence which is illustrated in Figure 3.

```
for item in window:
   if index of aspect current item = index of aspect
   next item AND index of sentence current item =
   index of sentence next item:
      remove next item from window
```

Fig. 3 Matrix Row Elimination

Before checking the sentence structure, the initial index and final index for each sentence segment are determined according to the location of the keyword (window). The initial index of sentence segments is determined using steps in Figure 4. The beginning of the sentence segment is determined based on the punctuation ('PUN' tag) and the conjunction ('CCN' tag) in the sentence. If both are not found, the beginning of the sentence is determined according to the location of the window in the sentence (beginning, middle, or end of sentence).

Steps in Figure 5 do determination of the end of the sentence segment. The end of the sentence segment is also determined based on the 'PUN' tag and the 'CCN' tag found in the sentence. If neither is found, a negation and adjective are checked, which is possible at the beginning of a sentence segment.

The negation word was chosen to be the end segment of a sentence only if the previous segment has already formed a clause. While the adjective is chosen as one of the candidates for the end of the current sentence segment, which is part of the next sentence segment, only if the previous segment has formed a clause and adjectives are found more than one in the right window.

```
for item in window list:
  left window = []
  right window = []
  for i in range (0, index of current window):
    add postag[i] into left window
  for i in range (index of current window, length
  sentence):
    add postag[i] into right window

  if 'PUN' in left window:
    awal = index of last 'PUN'
  else if 'CCN' in left window and 'PUN' not in
  left window:
    awal = index of last 'CCN'
  else if ('PUN' not in left window and 'CCN' not in
  left window and 'NEG' not in left window) or
  ('NEG' in left window and (distance of current
  window and 'NEG' > 2 OR distance of previous
  window and 'NEG' <= 2))
    if window at the beginning of sentence:
      awal = 0
    else if window at the middle of sentence:
      awal = akhir of previous item
    else:
      awal = indeks of current window
```

Fig. 4 Determination of the Initial Index of Sentence Segments

```
if 'NEG' in right window:
  tags_cek = []
  cek = index of last 'NEG'
  for i in range(awal, cek):
    add postag[i] into tags_cek
if 'ADJ' in right window:
  tags_cek = []
  cek = index of last 'ADJ'
  for i in range(awal, cek):
    add postag[i] into tags_cek

if 'PUN' in right window:
  akhir = index of last 'PUN'
else if 'CCN' in right window AND 'PUN' not in right
window:
  akhir = index of last 'CCN'
else if 'NEG' in right window AND 'CCN' not in right
window AND 'PUN' not in right window AND ((postag
current window = 'NNO' AND ('ADJ' in tags_cek OR
'VB' in tags_cek)) OR (postag current window = 'VB'
AND ('NNO' in tags_cek OR 'ADJ' in tags_cek)) OR
(postag current window = 'ADJ' AND ('NNO' in
tags_cek OR 'VB' in tags_cek)):
  akhir = cek
else if 'ADJ' in right window AND 'CCN' not in right
window AND 'PUN' not in right window AND ('NEG'
not in roght window or ('NEG' in right window and
(frequency of 'ADJ' in tags_cek > 1 OR 'ADJ' in left
window))) AND (postag current window = 'NNO' AND
('ADJ' in tags_cek OR 'VB' in tags_cek)) OR (postag
current window = 'VB' AND ('NNO' in tags_cek OR
'ADJ' in tags_cek)) OR (postag current window = 'ADJ'
AND ('NNO' in tags_cek OR 'VB' in tags_cek)):
  akhir = cek
else:
  if window at the end of sentence:
    akhir = length sentence
  else:
    akhir = index of next window
```

Fig. 5 Determination of the End Index of Sentence Segments

After the initial and final index of sentence segments are obtained, the sentence structure is checked with steps in Figure 6. The distance between the window and the words that precede and follow the window is chosen based on the review data pattern used by the user in providing his review. This distance can differ according to the data domain used.

The last step is sentences segmentation. At this step, the new starting point and endpoint are determined again according to the previously created rule shown in Figure 4 and Figure 5. Then the tokens are joined from the start index to the end index by separating them with whitespace.

```
if (postag = 'NNO' and ((followed or preceded by 'VB'
and right window <= 3 and left window <= 3) or
(follow or preceded by 'ADJ' and right window <= 5
and left window <= 3) or (follow or preceded by =
'KUA' and right window <= 3 and left window <= 2)))
or (postag = 'VB' and ((follow or preceded by = 'NNO'
and right window <= 3 and left window <= 3) or
(follow or preceded by = ADJ and right window <= 5
and left window <= 2))) or (postag = ADJ and ((follow
or preceded by = 'NNO' and right window <= 3 and
left window <= 5) or (follow or preceded by = 'VB' and
right window <= 2 and left window <= 5))):
  continue
else:
  if aspect in sentence = 1:
    remove current item from window list
  else:
    if window at the beginning of sentence:
      remove next item from window list
    else:
      remove current item from window list
```

Fig. 6 Sentence Structure Checking

D. *Sentiment Classification Using Support Vector Machine*

The steps taken in the sentiment classification process include:

1) Sentiment Labelling that groups data into three categories. The categories are positive, negative, and neutral. The amount of sentiment labeling the data, which will then be used as learning data to determine sentiment polarity, is shown in Table 3.

2) Stopwords Removal, stopwords are deleted again to delete common words that are repeatedly used in reviews but do not determine the sentiment polarity of the sentence and only use important information to be included in the classification process.

3) TF-IDF Weighting is performed on the data generated by the previous process. Using the TF-IDF method, weighting is done with steps calculate term frequency (TF), calculate inverse document frequency (IDF), and calculate the TF-IDF using Equation 1. IDF is calculated using Equation 2. This process uses Tfidfvectorizer() from Python.

$$Wdt = tf\ dt * IDF\ t \tag{1}$$

$$IDF = log\left({}^{D}/_{df}\right) \tag{2}$$

4) Sentiment Classification will be done using the Support Vector Machine Method (SVM), a technique for making predictions, both in classification and regression [22]. SVM has the basic principle of linear classifier, but SVM has been developed to work on non-linear problems by incorporating the concept of the kernel in a high-dimensional workspace. To get the best hyperplane, the same as maximizing the margin or distance between two sets of objects from different classes as in Figure 7. This study uses a linear SVM python library that implements the multi-class SVM method one against all.

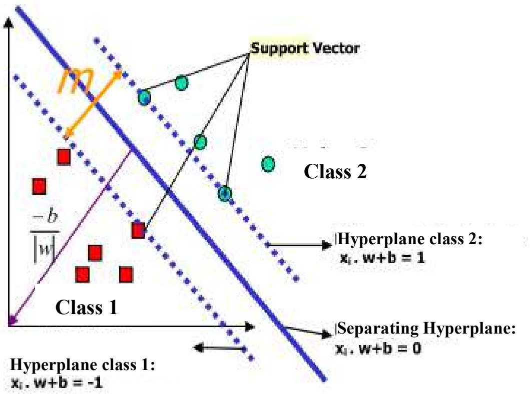

Fig. 7 The Best Hyperplane with The Biggest Margin[22]

E. Sentences Selection and Summary Generation

The method commonly used in selecting sentences from a document is based on the weight of the sentences sorted in descending order [23]. Before we calculate the weight of the sentence, we join the sentences that have the same aspect, polarity, and reviewer to gather all information about an aspect from one reviewer. The features used in this study are sentence to sentence similarity and sentences length.

Sentence to sentence similarity is the similarity between one sentence and another sentence. Sentence similarity can be calculated using a cosine similarity measure. The weight for this feature is obtained by calculating the ratio of the number of similarities of S sentences to other sentences and the maximum number of similarities shown in Equation 3 while similarity is calculated using Equation 4 [24].

$$\text{Score } S_i = \frac{\text{Sum of Sentences Similarity in } S_i}{\text{Max(Sum of Sentences Similarity)}} \quad (3)$$

Where,

$$\text{Similarity} = \cos(d \cdot q) = \frac{d \cdot q}{\|d\|\|q\|} = \frac{\sum_{i=1}^{|v|} d_i q_i}{\sqrt{\sum_{i=1}^{|v|} d_i^2}\sqrt{\sum_{i=1}^{|v|} q_i^2}} \quad (4)$$

Sentence length is the number of words in a sentence, where this feature serves to filter out sentences that are too short. Short sentences are not desired in summary. To measure the importance of sentences based on sentence length, normalization of sentence length is used, which is the ratio of words that appear in a sentence to the longest number of words in a sentence shown in Equation 5 [24].

$$\text{Score } S_i = \frac{\text{No. word occuring in } S_i}{\text{No.word in longest sentences}} \quad (5)$$

Sentence weight is obtained by calculating the similarity between one sentence with other sentences that have been grouped into the proper aspects and polarity using Equation 3 called F1-weight. This is intended to get the sentence that has the highest similarity (most often given in comments). Then, filter the sentence based on the F1-weight with a threshold of 0.8. Finally, calculate the weight based on the sentence length feature (F2-weight) to determine a sentence with the most information using Equation 5.

F. Summary Evaluation

After the final summary is obtained, an evaluation of the summary is carried out using the human environment. Summary evaluation is done by giving statements related to the parameters used to measure summary results. The statement given are shown in Table 1. The human evaluators will give a score to this statement. A five-point scoring scale from 1 to 5 represents a strong disagreement to a strong agreement.

TABLE I
STATEMENT LIST FOR RESPONDENTS

#	Statements
Q1.	The sentences can be understood clearly according to their meaning.
Q2.	Summary does not contain overlapping or redundant information about the product.
Q3.	The combination of sentence segments, in summary, can be received or understood.
Q4.	The sentences are suitable with their aspect category.
Q5.	The sentences are suitable with their sentiment or polarity.
Q6.	The sentences chosen and displayed in summary can represent the whole opinion in the reviews.
Q7.	The summary can help in finding specific information about the product.
Q8.	The summary can help determine whether to buy the product or not.
Q9.	The summary saves time in finding the desired information about the product.

III. RESULTS AND DISCUSSION

A. Aspect and Keyword Selection

After a manual inspection of the frequency of each class of words that appear, aspects and keywords are chosen as in Table 2.

TABLE II
ASPECTS AND KEYWORDS

Aspect	Keywords
kandungan	**Noun:** *alkohol, kandung, bahan, kadar* **Verb:** *kandung* **Adjective:** *aman, organik*
efek	**Noun:** *jerawat, efek, hasil, pecah, masalah, merah, iritasi, ubah, manfaat, fungsi, beruntus, bruntusan, komedo, guna, putih, reaksi, pengaruh, bintik,hitam* **Verb:** *beruntus, jerawat, cahaya, pengaruh, guna, merah, cerah, kering, bersih, pecah* **Adjective:** *kering, halus, lembut, kusam, mulus, bersih, pecah, ampuh, sehat, signifikan, kencang, cahaya, irita, cerah si, merah*
kemasan	**Noun:** *toples, kemas, botol, wadah* **Verb:** - **Adjective:** *higienis, ribet, tumpah*
aroma	**Noun:** *wangi, bau, aroma, parfum* **Verb:** *sengat* **Adjective:** *sengat, wangi, strong, bau*
sensasi	**Noun:** *sensasi, minyak, rasa, panas, lembab, gatal, cekit* **Verb:** *rasa, lembab, tenang, minyak, dingin, adem, segar* **Adjective:** *dingin, lembab, panas, lengket, segar, kenyal, minyak, tenang, perih, gatal, nyaman, sejuk, licin, ringan*
harga	**Noun:** *harga, kantong* **Verb:** *jangkau* **Adjective:** *murah, pantas, mahal, jangkau, hemat*
tekstur	**Noun:** *tekstur* **Verb:** - **Adjective:** *cair, kental*
daya serap	**Noun:** - **Verb:** *serap, resap* **Adjective:** -
isi	**Noun:** *isi* **Verb:** - **Adjective:** -

B. Sentences Segmentation Based Aspect

Sentence segmentation is done by matching keywords with words in sentences, determining the initial and final index, checking sentence structure, determining the new start and ending indexes, and merging word tokens into a sentence. Segmentation or separation of sentences is carried out by considering the POS-tag of each word and the position or distance of the word from the keyword (window) with supporting words that precede (left-window) or follow it (right-window). As an example, is a sentence in Figure 8 will produce segmentation result in Table 3

Harga/NNO nya/PRK **jangkau/VBP isi/NNO** nya/PRK banyak/KUA awet/ADJ sekali/ADV ,/PUN ketika/CSN di/PPO aplikasi/VBT ke/PPO wajah/NNO cepat/ADJ **serap/VBT** ke/PPO kulit/NNO tidak/NEG buat/VBT **lengket/ADJ** sama/ADJ sekali/ADV buat/VBT kulit/NNO jadi/VBI haluss/NNO ,/PUN **merah**/NNO diwajah/VBP jadi/VBI **tenang**/VBT ,/PUN no/NNO **aroma/NNO** ./PUN seluruh/ADJ bagus/ADJ

Fig. 8 Lemma and POS-Tags of Review

TABLE III
RESULTS OF SENTENCE SEGMENTATION

Sentence Segment	U	S	T	Aspect
harga nya terjangkau	0	0	0	*harga*
isi nya banyak awet sekali	0	0	3	*isi*
ketika di mengaplikasikan ke wajah cepat menyerap ke kulit	0	0	9	*daya serap*
tidak membuat lengket sama sekali membuat kulit jadi haluss	0	0	19	*sensasi*
Kemerahan diwajah jadi menenangkan , no aroma	0	0	29	*efek*
keseluruhan bagus	This is not included because it does not contain aspects.			

C. Sentiment Classification Using Support Vector Machine

In the sentiment classification stage, the first step that must be done is to divide the data into training data and test data to test the Support Vector Machine method. The dataset used is the result of grouping sentences according to aspects in detail in Table 4, which is then divided by a ratio of 80% used for training and 20% for testing data. The classification implementation uses the linear SVM.

After testing, the confusion matrix table is obtained as in Table 5. Based on the table, the accuracy obtained is 75%. As supporting information, we also perform 10-folds cross-validation by dividing the data into 10 folds that each fold is used as a learning and testing set at some point. The average accuracy obtained is 74%.

TABLE IV
DETAILS OF TRAINING DATA AND TEST DATA

Class	Number of Training Data	Number of Test Data
Positive	2,487	613
Negative	1,123	273
Neutral	506	144

TABLE V
CONFUSION MATRIX

Actual	Prediction		
	Negative	Neutral	Positive
Negative	180	10	83
Neutral	34	63	47
Positive	54	27	532

D. Sentences Weighting and Selection

For example, sentence selection of 5 sentences from the "effect-positive" category was made from the word (bow) bag in Table 6. In Table 6, data numbers 1 and 3 are not considered because the value of F1 weight is below the threshold. After data is filtered based on the F1 weight value, the F2 weight calculation is performed on the filtered data.

TABLE VI
SENTENCES WEIGHT

#	Bag of Words (BoW)	F1	F2
1	*merah wajah tenang no aroma*	0.782	
2	*terus bilas pakai air cerah*	0.892	0.263
3	*pas coba pakai aloe gel merk lokal bruntusan kurang sekali*	1	0.526
4	*tidak pecah tidak whitehead tambah*	0.782	
5	*sampel innisfree coba harap hasil bagus jam olesin aloe jerawat langsung kempis allag tangis langsung counter bel penuh ukur*	0.889	1

After weighting feature F2 is obtained, the sentence will be returned to the previous form, where the highest weight is displayed as a summary. As seen in Table 6, the sentence to be chosen as a summary is the fifth sentence because it has the highest weight. The original sentence from the data was "*ada sampel dari innisfree, saya mencoba dengan harapan hasil bagus; beberapa jam setelah saya olesin aloe vera ini itu jerawat langsung kempis iya allag mau menangis langsung saya ke counter untuk membeli penuh ukuran*". A summary obtained from all the reviews of Innisfree Aloe Revital Gel is shown in Table 7. The number beside positive and negative shows how many reviews or people have expressed about an aspect with a certain polarity.

TABLE VII
SUMMARY OF THE REVIEWS

INNISFREE ALOE REVITAL SOOTHING GEL

EFEK

Positive (232):
dipakai ke rambut setiap habis keramas begitu rambut nya kering jadi halus sekali rambut nya , dipakai ke wajah setiap hari juga baik-baik saja malah kalau pas lagi ada jerawat saya pakai ini yang banyak pas di daerah jerawat tidak sampai 3harian sudah kempis; sudah tidak pernah lagi mencetin jerawat semenjak punya ini karena begitu muncul langsung totol ini kempis

Negative (121):
terlihat sekali hasil nya komedo ku banyak berkurang dan beruntusan ku berkurang sekali i dont know if product klaim tapi di karena setau saya ini lebih ke pelembab , minyak control; terus saya juga make produk lain, dan kayak nya membersihkan di kulit saya berdoa semoga hanya membersihkan bukan tidak cocok beruntusan ku tumbuh di tempat yang saya tidak pernah beruntusan sebelumnya daerah pinggir mulut sampai pinggir hidung ke dekat pipi sudah saya coba berhenti produk nya untuk menghilangkan dulu beruntusan nya tapi produk ini tidak bisa bantu untuk menghilangkan itu , tapi on a bagus side , tiap ada jerawat mau muncul saya pakai ini semaleman terus tidak jadi muncul , atau maksimal make 2 hari tidak jadi mateng , dan untuk kulit saya yang kombinasi.

KANDUNGAN
Positive (132):
alasan saya lebih memilih mencoba produk ini dibanding yang natrep karena ini tidak ada alkohol , paraben , dan tidak pakai pewangi buatan begitu jadi aman untuk kulit sensitif
Negative (65):
tidak mengandung alkohol; tidak mengandung aroma, jadi cocok untuk kulit yang sensitif
KEMASAN
Positive (52):
aloe revital soothing gel nya innisfree ini yang paling saya suka kemasan nya yang botol; lebih higienis saja jadi nya tidak perlu colek
Negative (92):
kemasan nya botol jadi tetap higienis kalau ambil isi nya; sayang nya kemasan nya terlalu besar untuk dibawa , saran kalau mau dibawa kalian bisa pindahin ke tempat toples botol lain yang lebih mini
AROMA
Positive (68):
keseluruhan tidak ada beda nya hanya innsifree tidak ada wangi nya sama sekali
Negative (24):
tidak ada aroma parfum nya
SENSASI
Positive (252):
pas pertama kali pakai duh sensasi dingin nya itu terasa tenang sekali kaleem sekali; dan saya suka sekali pakai dia untuk jadi masker tidur jadi pas bangun tidur begitu muka rasanya kenyal sekali; kadang kalau pulang ngampus pas panas habis bersihin make uo cuci muka langsung pakai produk ini langsung dingin cinta kali
Negative (47):
pertama mencoba aloe vera gel itu yang merk nr tapi saya tidak merasakan efek yang signifikan bahkan seperti tidak memberi efek apa apa paling hanya sugesti saja kalau pakai ini jadi lembab; dan effect nya muka jadi benaran lembab sekali terus seperti terlihat begitu di kaca muka saya jadi terlihat lebih
HARGA
Positive (73):
harga nya juga terjangkau sekali , di grand indonesia membeli nya dengan harga 100 ribu isi nya 300 ml setengah tahun juga tidak habis
Negative (8):
sayang nya harga nya menurut ku mahal kkekwk membeli lagi iya soal nya tidak cocok terima kasih yang sudah baca
TEKSTUR
Positive (32):
tekstur nya enak tidak lengket menurut ku pas tekstur nya gel
Negative (2):
karena tekstur nya gel jadi pas awal make malah pas kering
DAYA SERAP
Positive (151):
yang paling saya notice dari brand ini dibanding product aloe yang lain adalah dia cepat sekali menyerap nya; walaupun dia cepat menyerap nya
Negative (10):
menyerap nya sedikit lebih lama dari natrep; lama nunggi menyerap
ISI
Positive (31):
ini senang sekali sudah mau 6 bulan belum habis juga isi nya; isi nya benar-benar banyak
Negative (1):
ternyata di saya malah tidak cocok membuat jidat ku beruntusan sedih mana isi nya banyak sekali

E. Experiment and Evaluation

We try to summarize the review data of Product-3 (new data as testing) with 462 reviews. After sentence segmentation-based aspect using Rule-Based, 1763 sentences were obtained according to its aspects. The confusion matrix obtained from the aspect categorization process is shown in Table 8. From this confusion matrix, the accuracy score obtained is 92%. Misclassification is most commonly found in "*kemasan*" and "*efek*" classes that should not be included. This is because of the keyword "*botol*" in the "*kemasan*" aspect used by reviewers to discuss how many bottles they have spent. Examples are in the sentence "*ini sudah botol ke 2 berangkat ke 3*". While the "effect" class error enters the "sensation" class because the reviewer uses the word "*rasa*", but the description that follows enters into the "*efek*" aspect. As an example, is in the sentence "*terus tidak tahu kenapa setelah seminggu pemakaian saya merasa pori-pori di wajah ikut kecil juga*".

TABLE VIII

CONFUSION MATRIX OF ASPECT CATEGORIZATION

	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	93	1	1	0	0	0	1	1	2	0
1	0	163	1	0	0	0	0	0	1	0
2	0	3	461	0	0	1	1	4	15	2
3	0	1	40	0	2	0	3	15	10	0
4	0	0	1	0	81	3	1	1	0	0
5	0	0	0	0	0	29	0	3	0	0
6	1	0	1	0	0	0	215	0	1	0
7	2	0	1	0	2	0	2	167	0	0
8	1	1	5	0	0	0	2	3	370	3
9	0	0	0	0	0	0	0	0	0	46

Horizontal: actual,

Vertical: prediction, 0: *aroma*, 1: *daya serap*, 2: *efek*, 3: general, 4: *harga*, 5: *isi*, 6: *kandungan*, 7: *kemasan*, 8: *sensasi*, 9: *tekstur*

TABLE IX

CONFUSION MATRIX OF SENTIMENT CLASSIFICATION

Actual	**Prediction**		
	Negative	**Neutral**	**Positive**
Negative	156	26	102
Neutral	37	68	70
Positive	217	55	1,033

Sentences are then classified based on their polarity using Support Vector Machine. The confusion matrix obtained from the sentiment classification process is shown in Table 9. From this confusion matrix, the accuracy score that obtained is 71,2%. Sentiment classification errors are mostly found in the class "positive" who enter the class "negative". Examples of errors are found in the sentence "*kemasan nya botol jadi tetap higienis kalau ambil isi nya*". "*higienis*" has the highest number of frequencies in the "negative" class because many data samples use the phrase "*kurang higienis*" or "*tidak higienis*" while in this study, we do not handle negation words that can change the meaning of a word. In the case of misclassification, "negative" to "positive" also affects because of the same thing. As an example, is in the sentence "*memang kalau dipakai tidak ada sensasi dingin seperti merek sebelah*" there is the word "*dingin*" which is most commonly found in the "positive" class.

With the accuracy that has been obtained, we surveyed the respondent about the summary. Respondents were asked to rate how the summary that has been produced in this study using the questions in Table 1. Respondents were presented with the original review text of 462 reviews of the Innisfree Aloe Revital Soothing Gel product, then the summary results obtained are shown in Table 7. This summary evaluation involved six respondents concerned with skin care products and always read comments or reviews before buying a skin care product. The evaluation results are shown in Table 10. From the table, we get the average score from 3 to 4. The word “SA” means they strongly agree with the statement, or each point's result is very good. “A” means they agree with the statement or the result in each point is good.

TABLE X
HUMAN EVALUATION RESULT

R	Question Score								
	1	2	3	4	5	6	7	8	9
R1	4	4	3	3	4	4	5	4	4
R2	5	3	5	3	3	4	5	5	5
R3	4	4	4	3	3	4	4	4	3
R4	4	3	4	4	3	3	4	4	4
R5	4	4	4	4	4	4	4	4	4
R6	4	3	4	4	3	3	4	4	4
R7	4	4	4	4	4	4	4	5	5
R8	4	4	4	4	5	4	5	4	4
R9	4	3	2	4	3	4	4	5	4
R10	5	5	5	4	4	4	5	5	5
R11	4	2	4	3	4	4	2	4	5
R12	5	4	4	4	3	3	5	4	4
R13	4	4	4	4	4	4	3	3	3
R14	4	4	5	4	4	4	4	5	5
R15	4	4	4	4	4	3	4	4	4
%	84	73	80	75	73	75	83	85	84
Result	SA	A	SA	A	A	A	SA	SA	SA

IV. CONCLUSION

The Rule-Based Method is proposed to identify aspects in a sentence and to split a sentence into several segments. Based on our evaluation of its classification results to 1,763 sentences, its accuracy achieves 92%. This score should have increased if we chose our keywords more carefully by considering various cases that may occur. As a consequence, a keyword must not be a common word that may represent two or more aspects. Using SVM, the first sentiment classification was done to all sentences' segments by dividing them 80%:20% for train data and test data. The test results showed 75 % as its accuracy score. After all sentence segments were rejoined by aspect and polarity, then the second sentiment classification using SVM was executed to those "new sentences" and gained accuracy 71,2%. Those misclassifications occurred because we did not handle negation words that can change the meaning of a word.

The summary of this study is presented per aspect and per polarity. Each displayed sentence on the summary is the most representative sentence selected from some sentences with the same aspect and polarity. Based on the human evaluation, the summary has received an average score of 79.1% or 3.95 (on a scale from 1 to 5). They agree that the summary can help them save their time to find specific information about the product and determine whether to buy the product. For future work, we can consider the negation word to reduce misclassification in the sentiment classification step. Overall, the proposed methods have provided state of the art in aspect-based opinion summary in Indonesian.

REFERENCES

[1] R. E. López Condori and T. A. Salgueiro Pardo, "Opinion summarization methods: Comparing and extending extractive and abstractive approaches," *Expert Syst. Appl.*, vol. 78, no. C, pp. 124–134, 2017, doi: 10.1016/j.eswa.2017.02.006.

[2] C. Surber and J. Kottner, "Skin care products: What do they promise, what do they deliver," *J. Tissue Viability*, vol. 26, no. 1, pp. 29–36, 2017, doi: 10.1016/j.jtv.2016.03.006.

[3] T. A. Tran, J. Duangsuwan, and W. Wettayaprasit, "Automatic Aspect-Based Sentiment Summarization for Visual, Structured, and Textual Summaries," *ECTI Trans. Comput. Inf. Technol.*, vol. 15, no. 1, pp. 50–72, 2021, doi: 10.37936/ecti-cit.2021151.237565.

[4] M. E. Moussa, E. H. Mohamed, and M. H. Haggag, "A survey on opinion summarization techniques for social media," *Futur. Comput. Informatics J.*, vol. 3, no. 1, pp. 82–109, 2018, doi: 10.1016/j.fcij.2017.12.002.

[5] C. F. Tsai, K. Chen, Y. H. Hu, and W. K. Chen, "Improving text summarization of online hotel reviews with review helpfulness and sentiment," *Tour. Manag.*, vol. 80, no. April, p. 104122, 2020, doi: 10.1016/j.tourman.2020.104122.

[6] S. Xiong, K. Wang, D. Ji, and B. Wang, "A short text sentiment-topic model for product reviews," *Neurocomputing*, vol. 297, pp. 94–102, 2018, doi: 10.1016/j.neucom.2018.02.034.

[7] R. Mukherjee, P. Goyal, H. C. Peruri, S. Bhattacharya, U. Vishnu, and N. Ganguly, "Read what you need: Controllable Aspect-based Opinion Summarization of Tourist Reviews," *arXiv*, pp. 1825–1828, 2020.

[8] W. M. Wang, Z. Li, Z. G. Tian, J. W. Wang, and M. N. Cheng, "Extracting and summarizing affective features and responses from online product descriptions and reviews: A Kansei text mining approach," *Eng. Appl. Artif. Intell.*, vol. 73, no. October 2017, pp. 149–162, 2018, doi: 10.1016/j.engappai.2018.05.005.

[9] I. El Alaoui, Y. Gahi, R. Messoussi, Y. Chaabi, A. Todoskoff, and A. Kobi, "A novel adaptable approach for sentiment analysis on big social data," *J. Big Data*, vol. 5, no. 1, p. 12, 2018, doi: 10.1186/s40537-018-0120-0.

[10] S. Lamsiyah, A. El Mahdaouy, B. Espinasse, and S. El Alaoui Ouatik, "An unsupervised method for extractive multi-document summarization based on centroid approach and sentence embeddings," *Expert Syst. Appl.*, vol. 167, no. September 2020, p. 114152, 2021, doi: 10.1016/j.eswa.2020.114152.

[11] S. K. D'Mello and J. Kory, "A review and meta-analysis of multimodal affect detection systems," *ACM Comput. Surv.*, vol. 47, no. 3, pp. 31–38, 2015, doi: 10.1145/2682899.

[12] M. R. Ramadhan, S. N. Endah, and A. B. J. Mantau, "Implementation of Textrank Algorithm in Product Review Summarization," *ICICoS 2020 - Proceeding 4th Int. Conf. Informatics Comput. Sci.*, 2020, doi: 10.1109/ICICoS51170.2020.9299005.

[13] K. Yauris and M. L. Khodra, "Aspect-based summarization for game review using double propagation," *Proc. - 2017 Int. Conf. Adv. Informatics Concepts, Theory Appl. ICAICTA 2017*, pp. 0–5, 2017, doi: 10.1109/ICAICTA.2017.8090997.

[14] Derisma, D. Yendri, and M. Silvana, "Comparing the classification methods of sentiment analysis on a public figure on indonesian-language social media," *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 8, pp. 1214–1220, 2020.

[15] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.

[16] E. Sutoyo and A. Almaarif, "Twitter sentiment analysis of the relocation of Indonesia's capital city," *Bull. Electr. Eng. Informatics*, vol. 9, no. 4, pp. 1620–1630, 2020, doi: 10.11591/eei.v9i4.2352.

[17] G. A. Buntoro, R. Arifin, G. N. Syaifuddiin, A. Selamat, O. Krejcar, and H. Fujita, "Implementation of a Machine Learning Algorithm for Sentiment Analysis of Indonesia's 2019 Presidential Election," *IIUM Eng. J.*, vol. 22, no. 1, pp. 78–92, 2021, doi: 10.31436/IIUMEJ.V22I1.1532.

[18] F. Binsar and T. Mauritsius, "Mining of Social Media on Covid-19 Big Data Infodemic in Indonesia," *J. Comput. Sci.*, vol. 16, no. 11, pp. 1598–1609, 2020, doi: 10.3844/JCSSP.2020.1598.1609.

[19] I. G. M. Darmawiguna, G. A. Pradnyana, and I. B. Jyotisananda, "Indonesian sentiment summarization for lecturer learning evaluation by using textrank algorithm," *J. Phys. Conf. Ser. Pap.*, vol. 1810, no. 012024, 2021, doi: 10.1088/1742-6596/1810/1/012024.

[20] D. Ekawati and M. L. Khodra, "Aspect-based sentiment analysis for Indonesian restaurant reviews," in *Proceedings - 2017 International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2017*, 2017, pp. 1–6, doi: 10.1109/ICAICTA.2017.8090963.

[21] N. A. Salsabila, Y. Ardhito, W. Ali, A. Septiandri, and A. Jamal, "Colloquial Indonesian Lexicon," 2018.

[22] B. Santosa, "Data mining teknik pemanfaatan data untuk keperluan bisnis," *Yogyakarta Graha Ilmu*, vol. 978, no. 979, p. 756, 2007.

[23] X. Mao, H. Yang, S. Huang, Y. Liu, and R. Li, "Extractive summarization using supervised and unsupervised learning," *Expert Syst. Appl.*, vol. 133, pp. 173–181, 2019, doi: https://doi.org/10.1016/j.eswa.2019.05.011.

[24] S. A. Babar and P. D. Patil, "Improving performance of text summarization," *Procedia Comput. Sci.*, vol. 46, no. Icict 2014, pp. 354–363, 2015, doi: 10.1016/j.procs.2015.02.031.