**Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)**
Vol. 4, No. 1, March 2021, pp. 19 – 28
p-ISSN: 2614-3372 | e-ISSN: 2614-6150 

# Consumer Opinion Extraction Using Text Mining for Product Recommendations On E-Commerce

**[1]Erlina Halim, [2]Ronsen Purba, [3]Andri**
[1,2] Department of Informatics Engineering, STMIK Mikroskil
[3]Department of Information System, STMIK Mikroskil
Email: [1]erlina.halim@mikroskil.ac.id, [2]ronsen@mikroskil.ac.id , [3]andri@mikroskil.ac.id

**Article Info**
***Article history:***
Received Sep 25th, 2020
Revised Dec 03rd, 2020
Accepted Dec 17th, 2020

***Keyword:***
Text Mining
Lexicon
Classification
Consumer Opinion
E-commerce

**ABSTRACT**
Consumer opinion on e-commerce has a role in influencing consumers in the purchasing process and can be a recommendation. But the conditions of diverse opinions and the use of non-standard words are challenges in processing opinions. Opinion needs to be processed first by applying spelling normalization and repairs to the slang word using a slang dictionary that lists slang words and their conversions. The normalized opinion will be used to extract opinions using text mining with the lexicon approach. This approach requires a dictionary of words that contain opinions along with weight that are worth 1 to 5 for positive opinions and are worth -1 to -5 for negative opinions. The weight found for each existing opinion will be used to determine the classification. Classification uses the ratio of maximum positive opinion weight to the maximum weight of negative opinion. The resulting classification of opinions is positive, negative or neutral. The classification of opinion produced is positive, negative or neutral. Opinion classification is then compared with the rating classification to determine the level of accuracy. Comparison produces an accuracy of 80.34% by completing an opinion dictionary.


***Corresponding Author:***
Erlina Halim,
Department of Informatics Engineering,
STMIK Mikroskil,
Email: erlina.halim@mikroskil.ac.id

**DOI: http://dx.doi.org/10.24014/ijaidm.v4i1.10834**

## 1. INTRODUCTION

Consumer opinion extraction is the process of generating new and important information from opinions that contain statements about claims in services or products provided by e-commerce to consumers [1]. The information obtained will provide benefits for consumers and sellers. For sellers, this information will be used to manage customer satisfaction, trust, and loyalty for the long-term development of e-commerce [2]. For consumers, the information obtained can be a recommendation in the purchase process [3,4]. The selection of opinions based on the case there is a rating value that is not per under the opinion [5]. According to [6], the effect of opinion gives an influence on purchases by 73% - 87%. Automatic consumer opinion extraction is important [1]. Accurate information can be obtained by reading all opinions [7]. Besides, it is necessary to pay attention to language in opinions that do not pay attention to sentence structure, the use of non-formal language and the use of emoticons or images [8,9]. With diverse opinion conditions, it is necessary to apply text mining to change opinions understood by human language into opinions that can be understood by computers for further processing [10].

According to [11] studies, user behavior in the use of applications on Android IOS. The choice of using an application depends on the average rating given by other users. Problems that arise by only focusing on the rating are the discrepancy between rating and opinion and bias on the conclusion of user rating. The opinion is extracted using sentiment-lexicons and this approach can overcome the problem of data from various categories. The output of this research is in the form of theoretical studies and there are no measurements to calculate the efficiency of the resulting theory.

Research by [12], conducted an analytical study on the use of the lexicon in Indonesian language data. The approach taken in this study is the conversion of Indonesian into English, then applying the lexicon technique to the English data dictionary to get sentiment scores. User opinion data is obtained from the Google Play Store and App Store. The conversion process uses a Bing translator. Researchers assume the translate results are good enough and do not do the correction process manually. The results of this study indicate an average accuracy value of 0.68 for the classification of opinions of three types of sentiments, namely positive, negative and neutral. Some of the issues discussed in this study are the problem of data normalization, non-standard language structure, and ambiguous words.

Research by [13], have compared the lexicon technique with machine learning for sentiment analysis on social media. The data used are in the form of big data from 83 brands on Facebook with 850 samples of commentary data sourced from users. The data is used to determine positive or negative sentiment from consumers. The results of these studies indicate that both approaches have a similar level of accuracy. Also, both techniques can provide higher accuracy for detecting positive sentiment than negative sentiment.

Research by [14], extracted service opinion definite fitting Pertamina with a total of 150 data opinions using the lexicon approach. The dictionary of words used in this approach includes positive words, negative words, negation words, emoticons, and Indonesian dictionaries. In the positive and negative word dictionary, there are no weights, only the emoticon dictionary has positive (1) or negative (-1) weights. The determination of opinion classification is done by calculating the probability of the emergence of more dominant positive and negative words. In this research, the preprocessing applied is only normalization and tokenization. In comparing the extraction results with opinions, the researcher carries out the labeling process manually into positive, negative or neutral classes which are then compared with the results of opinion classification. The results of this study produce an accuracy of 0.66.

Based on the above research description, rating selection refers to Islamic research which finds that there is a mismatch between ratings and opinions. So that opinion extraction using lexicon can be considered to obtain a more accurate assessment of the product or service. This is supported by studies Dhaoui, et al who compared the results of accuracy lexicon with machine learning. Both methods produce the same accurate high to detect positive sentiment than negative sentiment. The Putri and Pamungkas study also used lexicon to process opinions with an average accuracy of 0.68. Accuracy results are obtained without applying data normalization, checking non-standard language structures and ambiguous words. So that this research will discuss issues that have not been resolved. This study also renewed Nurfalah and Suryani's research, which is a dictionary using a dictionary in the research of Putri and Pamungkas, which was translated into Indonesian, then added the preprocessing stages, namely stemming and stopword removal. By using the lexicon approach, two pieces of information can be obtained, namely the weight of opinion and classification. Opinion weights will be calculated to get a classification that will classify opinions into positive, negative or neutral classes. The results of the opinion classification will be compared with the rating that has been classified to get the level of accuracy.

## 2. RESEARCH METHOD

### 2.1. Classification

A very important part of data mining is the classification technique, which is how to study a set of data so that rules are produced that can classify or recognize new data that has never been studied. Classification can be defined as the process of declaring a data object as one of the pre-defined categories (classes). Classification is widely used in various applications, including fraud detection, customer management, medical diagnosis, sales predictions, and so on [15].

The classification model can be built based on the knowledge of an expert (expert). However, given the very large data set, the classification model is more often constructed using learning techniques in the field of machine learning. The learning process automatically on a data set can produce a classification model (target function) that maps data objects $x$ (input) to one of the classes $y$ that have been previously defined. So, the learning process requires input (input) in the form of a set of training data (training set) that is labeled (has class attributes) and outputs a form of a classification model [15].

Determination of classification aims to determine the sentiment of an opinion sentence. There are 3 ways to determine class, namely:

1. According to Nurfalah & Suryani [14], determination of sentiment is done by calculating the probability of the emergence of positive keywords and negative keywords. After knowing all the keywords and emoticons that have sentiment values, then the probability of the emergence of positive and negative sentiments which is more dominant is calculated. If the positive sentiment value is more dominant then the sentiment value for the sentence is positive, but if the negative sentiment value is more dominant then the sentiment value for the sentence is negative, but if the value is the same

between negative sentiment and positive sentiment then the sentiment value for the sentence is neutral. Formula 1 is a formula for determining sentiments:

$$sentiment\ value = \begin{cases} 1, & \sum sentiments + emotions > 0 \\ 0, & \sum sentiments + emotions = 0 \\ -1, & \sum sentiments + emotions < 0 \end{cases} \quad 1$$

2. According to Pamungkas and Putri [12], the determination of sentiments is done by calculating the number of positive (Spositive) keyword weights and the number of negatives (Snegative) keyword weights found in opinions. The formula for getting Spositive and Snegative can be seen in equation (2) and (3) below.

$$S_{positive} = \sum_{i \epsilon t}^{n} positive\ score_i \quad 2$$

$$S_{negative} = \sum_{i \epsilon t}^{n} |positive\ score_i| \quad 3$$

After the number of sentiments is obtained, Spositive and Snegative comparison is performed. If a positive value is greater than a negative value, then the sentence is a positive opinion. If the negative value is greater than the positive value, then the sentence is a negative opinion. If the positive and negative values are the same, then the sentence is a neutral opinion.

$$sentence_{sentiment} \begin{cases} positive\ \ if\ s_{positive} > s_{negative} \\ neutral\ \ if\ s_{positive} = s_{negative} \\ negative\ if\ s_{positive} < s_{negative} \end{cases} \quad 4$$

3. According to Thelwall [16], determination of sentiment is done by calculating the maximum weight of positive opinion (Spositive) and the maximum weight of negative opinion (Snegative) found in opinion. If Max (Spositive) is greater than Max (Snegative), then it is a positive class. If Max (Spositive) is smaller than Max (Snegative), then it is a negative class. If Max (Spositive) and Max (Snegative) are of the same value then they are neutral. Shows the classification formula 5.

$$sentence_{sentiment} \begin{cases} positive\ \ if\ max(s_{positive}) > max(|s_{negative}|) \\ neutral\ \ if\ max(s_{positive}) = max(|s_{negative}|) \\ negative\ if\ max(s_{positive}) < max(|s_{negative}|) \end{cases} \quad 5$$

In addition to classifications to determine opinion classes, classifications for ratings are also needed to be compared with opinions. Rating classification rules according to Xia and Jiang [17] can be seen in table 1 below.

**Table 1.** Rating Classification Rules [17]

| Rating | Kelas |
|---|---|
| 1-2 | Negative |
| 3 | Neutral |
| 4-5 | Positive |

Words of positive opinion are used to express desired conditions while negative words of opinion are used to express unwanted circumstances. Examples of positive opinion words are: beautiful, beautiful, good, and amazing. Examples of words of negative opinion are bad, poor, and terrible. Apart from individual words, there are also opinion phrases and idioms. Collectively, it is called the lexicon of opinion and is an instrument for sentiment analysis.

## 2.2. SentiStrength

SentiStrength [16] is a lexicon-based classification that exploits the sentiment lexicon that is built as well as combining entries from different linguistic resources. The SentiStrength algorithm is used to identify the polarity of sentiments in the text and detect the strength of the sentiments expressed. SentiStrength displays positive and negative sentiment scores, namely −1 (not negative) to −5 (very negative), 1 (not positive) to 5 (very positive) for each input text written in English. Based on research from psychology, revealed that the process of positive and negative sentiments runs in parallel (a mixture of emotions). Based on their algebraic numbers, SentiStrength can also report overall trinary scores, namely overall positive (score = 1), negative (score = −1) and neutral (score = 0).

The SentiStrength algorithm depends on the information used in the algorithm and can be adjusted. The information referred to as follows [16]:

1. Sentiment Dictionary (EmotionLookUpTable)
   Sentiment dictionary contains a collection of words that have been weighted with the power of sentiment 1 (has no positive sentiment) to 5 (has very strong positive sentiment), and -1 (has no negative sentiment) to -5 (has very strong negative sentiment). Sentiment dictionary is obtained from the translation of the English sentiment dictionary which has experienced the addition and subtraction of words based on observations in the process of developing this system.
2. Emoticon Dictionary (EmoticonLookUpTable)
   Emoticons are symbols or combinations of symbols that are usually used to describe human facial expressions that contain emotions or feelings in the form of messages or writing. For example, the symbol ":D" shows the expression of laughing emoticons. This data is generated manually based on observations in the process of developing this system. This emoticon dictionary is also given a weight that will determine the change in sentiment weight in a sentence.
3. Idiom Dictionary (IdiomLookupTable)
   Idiom Dictionary is an idiom that has a different meaning from the original word. For example, the word "slamming bone" does not mean slammed bone but rather a substitute expression for working hard. The Idiom Dictionary is also given a weight that will determine the change in sentiment weight in a sentence. This data is generated manually based on observations in the process of developing this system.
4. Slang words and Conversions (SlangLookupTable)
   Slang words are non-standard words that are often used, such as abbreviated words, commonly used social language and words that have writing errors.
5. BoosterWordList
   BoosterWordList is a word that can increase or decrease the intensity of word sentiments next to it. For example, the word "very happy" is more positive than the word "happy", the word "less happy" is no more positive than the word "happy". This word is given 1-2 weights to increase or decrease the word score beside it. This data is generated manually based on observations in the process of developing this system.
6. Negating Word (NegatingWordList)
   The word negation is a word contained in a sentence that can change the orientation of an opinion. For example, the word "naughty" is a word that means negative, but if the word "naughty" is preceded by the word negation "no" then the wording becomes "not naughty" which means positive. This word is saved in the form of a .txt file, not saved in the database. This data is generated manually based on observations in the process of developing this system.
7. Question Words (QuestionWords)
   The question word is a word contained in a sentence that can change the orientation of an opinion. For example, "are you angry?", Although there is the word "angry", this sentence has no positive or negative orientation, so it is classified as a neutral sentence. This word is saved in the form of a .txt file, not saved in the database. This data is generated manually based on observations in the process of developing this system.

To adjust SentiStrength for Indonesian, the first step is to replace the English glossary of Emotion-LookupTable.txt with Indonesian. Word lists are compiled by translating word lists in English and adding words that contain the most common sentiments of the analyzed opinion. Every word on EmotionLookupTable.txt must be marked with a sentiment score that shows the typical polarity and strength of the sentiment expressed using the following scheme:

[−5] Very strong negative sentiment
[−4] Strong negative sentiment
[−3] Moderate negative sentiment
[−2] Mild negative sentiment
[2] Mild positive sentiment
[3] Increase positive sentiment

[4] Strong positive sentiments
[5] Very strong positive sentiment

### 2.3. Confusion Matrix

The confusion matrix is useful for analyzing the quality of classifiers in recognizing tuples from existing classes. There are 4 tuples, namely True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). TP and TN state that the classifier recognizes tuples correctly, meaning that positive tuples are recognized as positive and negative tuples are recognized as negative. In contrast, FP and FN state that the classifier incorrectly recognizes tuples, positive tuples are recognized as negative and negative tuples are recognized as positive. The confusion matrix can be represented as a 2x2 matrix sown by table 2 [15].

Table 2. Confusion Matrix

| | Class | Positive | Negative |
|---|---|---|---|
| Prediction | Positive | True Positive (TP) | False Negative (FN) |
| | Negative | False Positive (FP) | True Negative (TN) |

Several general formulas can be used to calculate classification performance. The results of the accuracy (equation 6), precision (equation 7), and recall (equation 8) values are usually displayed as a percentage [15].

1. Accuracy
   Accuracy is the sum of the correct prediction proportions. The accuracy calculation formula.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

2. Precision
   Precision is the proportion of the number of relevant text documents identified among all the text documents selected by the system. Precision formula.

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

3. Recall
   Recall is the proportion of the number of relevant text documents identified among all relevant text documents in the collection. Recall formula.

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

Figure 1 will explain the conceptual framework of problem-solving from the research to be conducted. The e-commerce platform has grown rapidly with a variety of products and buying experiences with attractive offers. But the effect on consumer insecurity in recognizing the product to be purchased even though there is already a solution provided that is meta-data. Consumers still find it difficult to decide to buy just by knowing the product's features. So e-commerce provides a forum feature so that consumers get a picture of the product based on previous consumer experience and can provide an opinion on the product. However, with so many opinions available on the product, it is difficult for consumers to read all opinions and make decisions in buying products. If consumers only read a few opinions, then a biased view of the product can be generated [7]. To make it easier to translate consumer opinions, e-commerce utilizes ratings. The rating used is a star from a scale of 1 to 5. However, using a rating alone is not enough to be a reference to determine sentiment, because there are differences between ratings and opinions given [11].

The action that can be applied is implementing text mining to process all opinions by taking into account the opinion format [8]. There is a program to process consumer opinion in English with the name SentiStrength [16]. This program will be adopted to be able to process opinions in Indonesian. The workings of this program use the lexicon approach, which requires a dictionary of words in determining words that are opinions and are weighted. Based on the weight obtained in the opinion will be compared to the maximum weight between positive opinion and negative opinion to be categorized into positive, negative, or neutral classes. The classification produced for each opinion and grouped per item can provide benefits to both consumers and sellers. Consumers are helped in the buying process based on previous consumer experience. As for sellers, they can manage customer satisfaction, trust, and loyalty for the long-term development of e-commerce.

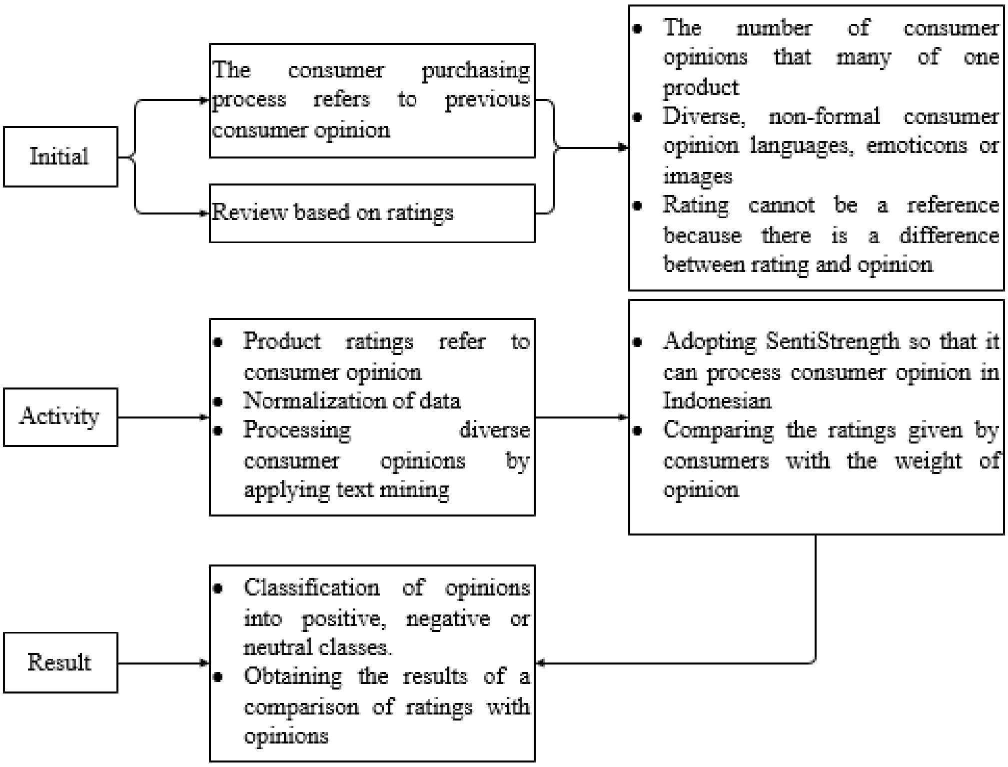


**Figure 1.** Research Conceptual Framework

The opinion class obtained will be compared with the rating class to determine the suitability of the rating value with the weight. In this study, the rating is assumed to be actual data with the consideration that rating is a value given directly by consumers. The classification of ratings refers to the rules put forward by Xia and Jiang [17], namely rating 1-2 including negative class, rating 3 including neutral class and rating 4-5 including positive class. After the rating class and opinion class are available, the comparison will be displayed in graphical form and confusion matrix. After obtaining a confusion matrix, the next step calculates accuracy, precision, and recall.

In this study, we solved three problems related to consumer extraction opinion for recommendation product: (1) many consumer opinions on e-commerce cause consumers to have trouble getting information, (2) opinion with a variety of forms and the use of non-standard words that can only be understood by humans, and (3) limited tools that can process consumer opinion in Indonesian. This research is an improvement to research by Pamungkas & Putri [12] and Nurfalah & Suryani [14]. The difference between the proposed research and previous research is shown in table 3.

**Table 3.** Proposed Research

| | Pamungkas dan Putri (2016) | Nurfalah dan Suryani (2017) | Proposed Research |
|---|---|---|---|
| Language in the Dataset | Indonesian conversion to English | Indonesian | Indonesian |
| Dictionary Language | English | Indonesian | Indonesian |
| Booster | yes | no | yes |
| Emoticon | yes | yes | no |
| Emotion | yes | yes | yes |
| Idiom | yes | no | yes |
| Negating | yes | yes | yes |
| Question | yes | no | yes |
| SlangLookup | yes | yes | yes |
| Preprocessing | | | |
| Remove Punctuatuion | no | no | yes |
| Case Folding | no | yes | yes |
| Spelling Normalization | no | yes | yes |
| Stopword Removal | no | no | yes |
| Stemming | no | no | yes |
| Tokenizing | yes | yes | no |
| Pos Tagging | yes | no | no |

| | Pamungkas dan Putri (2016) | Nurfalah dan Suryani (2017) | Proposed Research |
|---|---|---|---|
| Classification Determination | number of positive weights vs the number of negative weights | probability of the appearance of a positive word vs. a negative word | maximum positive weight vs. maximum negative weight |
| Classification Class | positive, neutral or negative | positive, neutral or negative | positive, neutral or negative |

This study uses primary data of 125,144 opinions collected online using crawling techniques from the Shopee e-commerce website in the "women's clothing" category. The selection of e-commerce and category refers to the "*katadata Indonesia*" survey in November 2018 which informs that Shopee occupies the first position while women's clothing ranks second [18]. The results of crawling in the form of the JSON (JavaScript Object Notation) file have many attributes that require the parsing process to get the attributes needed in this study and stored in a MySQL database. The parsing attribute is shown in table 4.

**Table 4.** Attributes of Parsing Result

| Field Name | Field Type | Description | Example |
|---|---|---|---|
| itemname | string | product name | Baju Rajut Boxy Premium |
| date | date | comment date | 2019-04-30 |
| starrate | int | rating | 5 |
| comment | string | comments given by the buyer | pengiriman cepet .. baju nya bagus .. sesuai harga .. admin nya ramah .. makasih shopee insaalloh jadi langganan nih. |

This research applies 3 processes to solve the raised issues, namely (1) the dictionary preparation process, (2) the opinion extraction process and (3) the comparison of the rating with the opinion. In preparing Indonesian dictionaries, the source used is the English Sentistrength dictionary. The translation process will be repeated for each dictionary and each line in the input file using the Google API provided by Han [19]. After the word is translated, the stemming process is carried out to get the basic word. Translations and weights will be saved in a new file.

After the dictionary is available, the stages of opinion extraction are as follows:

1. Dataset crawling assessment results from Shopee's e-commerce which is the consumer's response to the product or service provided by the seller. Reviews can be ratings and opinions in a variety of formats such as text, images or a combination of both.
2. Applying pre-processing to consumer opinion to change the format of opinion understood by human language to a format that can be understood by computers [19]. Several stages of text preprocessing that will be applied are remove punctuation, case folding, spelling normalization, stopword removal and stemming. After applying pre-processing, consumer opinion data is reduced by 538 because there are opinions that only contain punctuation or images.
3. To apply SentiStrength for Indonesian, the dictionary prepared in the previous stage will be read. Each word in the opinion will be matched with the word in the dictionary so that the weight is obtained.
4. The weight obtained will be calculated by determining the maximum weight for positive words and negative words. The determination of classification refers to the biggest weighting of the two words. If the positive word weight is greater then it will be classified into positive classes, if the negative word weight is greater, then to the negative class and if the values are the same then to the neutral class.

The results of the opinion classification will proceed to the testing stage by comparing the opinion class with its rating value. To be comparable, the class between opinion and rating must be the same, so ratings must be classified by the rules put forward by Xia and Jiang [17]. The results of the comparison will be displayed in graphical form and confusion matrix. The results of the confusion matrix are then used to calculate accuracy, precision, and recall.

## 3. RESULT AND ANALYSIS

Data analysis in this study applies descriptive statistics. This research data is in the form of text-based user opinions and ratings sourced from e-commerce websites. The word dictionary for the lexicon approach is sourced from the English-based SentiStrength dictionary. The dictionary from English will be converted to Indonesian using the Google API translator. After conversion, the manual entry will be added to support non-standard language structures such as slang, incomplete words, and so on. Distribution of data will be described using a histogram.

Testing the results of Indonesian-based opinion classification using text mining with the lexicon approach is done with a quantitative approach. The results of the opinion classification will be tested by comparing the results of the rating classification given by the user in the same comment. Evaluation of the model uses a confusion matrix to obtain the value of accuracy, precision, and recall. Testing is done for the dictionary of the results of conversion and also testing for the dictionary that has been converted plus a manual entry.

The following will be shown the process of weighting opinion starting from inputting opinion, pre-processing stage until the opinion class is obtained

1. Load Opinion in Indonesian
   "*Maaf kak,, yang ini bahannya tipis banget, trus juga kependekkan celananya,, lain kali diperbaiki yakk.. Terimakasih. Kualitas produk sangat baik. Produk original. Harga produk sangat baik. Kecepatan pengiriman sangat baik.*"
2. Pre-processing Result
   "*maaf kak ini bahan tipis banget terus pendek celana kali baik kualitas produk sangat baik produk harga produk sangat baik cepat kirim sangat baik*"
3. Weighting Opinion
   "*maaf* [2] *kak ini bahan tipis* [-2] *banget terus* [1] *pendek celana kali baik* [2] *kualitas produk sangat baik* [4] *produk harga* [2] *produk sangat baik* [4] *cepat kirim sangat baik* [4]"
4. Classification
   From the results of the weighting above, the maximum weight for positive sentiment: Max($\boldsymbol{S_{positive}}$) = 4 and the maximum weight for negative sentiment: Max($\boldsymbol{S_{negative}}$) = 2. Using the equation below, the class for the above opinion is positive.

$$Sentence_{sentiment} = \begin{cases} positive\ \ if\ Max(S_{positive}) > Max(|S_{negative}|) \\ neutral\ if\ Max(S_{positive}) = Max(|S_{negative}|) \\ negative\ if\ Max(S_{positive}) < Max(|S_{negative}|) \end{cases} \quad (9)$$

To improve extraction opinions, the conversion dictionary is also added to the entry manually based on the consumer opinions. The results of the distribution of opinion classes after the addition of entries can be seen in table 5 and diagram form in figure 2 below.

**Table 5.** Experiment Class Distribution

| Class | Count |
|---|---|
| Negative | 7.099 |
| Neutral | 21.895 |
| Positive | 95.612 |

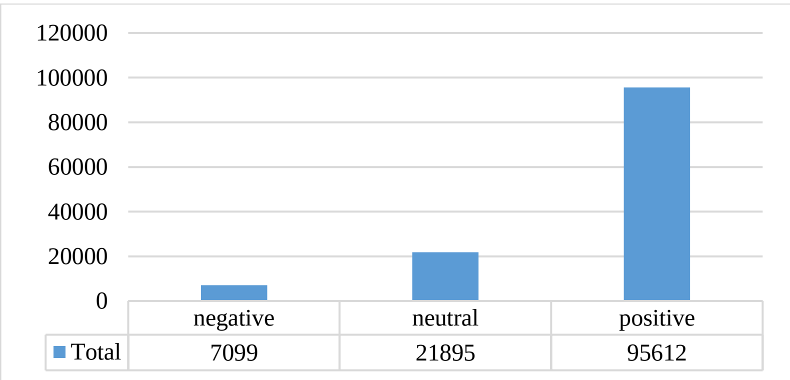


**Figure 2.** Experiment in Bar Diagram

The results of the extraction and classification of opinions in the previous section will be continued by comparing the number of rating class distributions with the opinion class in the experiment shown in figure 3.

Based on the graph above, there is a class difference between rating and opinion. The positive class on the rating is not entirely correct because the opinion contains negative words which cause the opinion to be included in the negative or neutral class. Then the confusion matrix for the experiment can be shown in table 6. The results of the accuracy in the experiment were 80.34%. For precision and recall for each class can be shown in table 7.

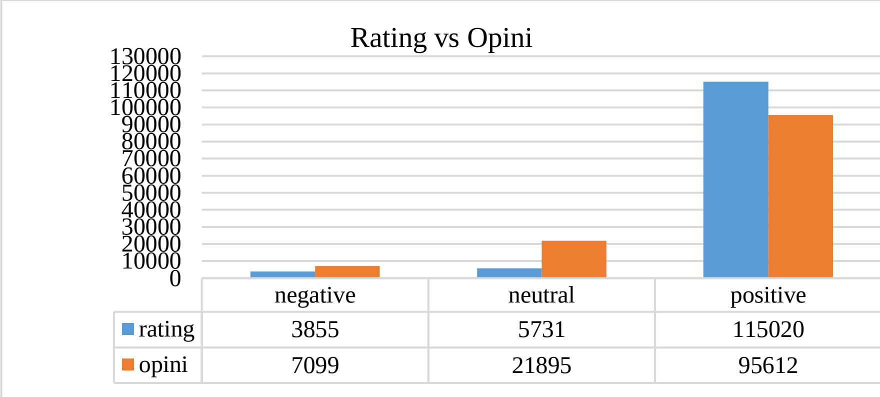


**Figure 3.** Graph Comparison of Rating Classes with Opinion Classes

**Table 6.** Confusion Matrix

| Class | Opinion | | |
|---|---|---|---|
| | Negative | Neutral | Positive |
| Negative | 2.379 | 1.348 | 3.372 |
| Neutral | 1.115 | 3.430 | 17.350 |
| Positive | 361 | 953 | 94.298 |

**Table 7.** Precision and Recall for the Experiment

| Class | Precision | Recall |
|---|---|---|
| Negative | 61.71% | 33.51% |
| Neutral | 59.85% | 15.67% |
| Positive | 81.98% | 98.63% |

The result of this experiment got better accuracy than similar experiment. The accuracy of this experiment got 80.34% better than experiment by Pamungkas produced 68% and experiment by Nurfalah and Putri produced 66%.

## 4. CONCLUSION

The results of the experiments show that words in the dictionary will effect accuracy of extraction opinion. The accuracy of 80.34% shows the suitability of the opinion class in detecting the correct class against the rating class of 124,606 opinions which is 100,107 right opinions and 24,499 wrong class opinions. Based on the results of research conducted, it can be concluded that: (1) opinion extraction using the lexicon approach to text mining can overcome the problem of translating many opinions on e-commerce and being able to classify each opinion into positive, negative, or neutral classes. (2) various consumer opinions and non-standard words can be normalized by applying spelling normalization and improvement of slang words by using a slang dictionary containing slang words and their conversions. (3) A comparison of rating classifications with opinion classifications of 124,606 data used produces an accuracy of 80.34%. (4) Accuracy can be improved by completing words of opinion in the word dictionary. Some suggestions for future research can be done by (1) Opinion extraction in this research results in the form of general opinion class. This research cannot find out opinions written referring to products, services or other information. (2) Completing opinions on dictionaries with the word synonym. (3) Opinions and ratings in this research are both experiencing the classification process into 3 classes so that it can be reviewed for the classification of opinions into 5 (five) classes so that it does not require a classification process for rating..

## REFERENCES

[1] Vikas, B. O. & Mungara, J. "An Enhanced Extraction and Summarization Technique with User Review Data for Product Recommendation to Customers", International Journal of Scientific Research in Science, Engineering and Technology, 2(6), 2016, pp. 25–30.

[2] Addepalli, S. L. et al. "A Proposed Framework for Measuring Customer Satisfaction and Product Recommendation for Ecommerce", International Journal of Computer Applications, 138(3), 2016, pp. 30–35.

[3] Ramanathan, V. & Meyyappan, T. "Twitter Text Mining for Sentiment Analysis on People's Feedback About Oman Tourism", 2019 4th MEC International Conference on Big Data and Smart City, ICBDSC 2019. IEEE, 2019, pp. 1–5.

[4] Sohail, S. S., Siddiqui, J. & Ali, R. "User Feedback Scoring and Evaluation of a Product Recommendation System", IEEE, 2014.
[5] Fan, Z., Chang, D. & Cui, J. "Algorithm in E-commerce Recommendation", 2018 5th International Conference on Industrial Economics System and Industrial Security Engineering (IEIS). IEEE, 2018, pp. 1–6.
[6] Tama, V. O., Sibaroni, Y. & Adiwijaya. "Labeling Analysis in the Classification of Product Review Sentiments by using Multinomial Naive Bayes Algorithm", Journal of Physics: Conference Series, 1192, 2019, p. 012036.
[7] Rajganesh, N., Nandhini, R. & Sumitha, M. "A Recommendation System for Online Products by Analyzing the Customer Feedback", International Journal of Computer Science Trends and Technology, 4(2), 2016, pp. 14–18.
[8] Al-Rubaiee, H. et al. "Techniques for Improving the Labelling Process of Sentiment Analysis in the Saudi Stock Market", International Journal of Advanced Computer Science and Applications, 9(3), 2018.
[9] Yassine, M. & Hajj, H. "A Framework for Emotion Mining from Text in Online Social Networks", Proceedings - IEEE International Conference on Data Mining, ICDM, December 2010, pp. 1136–1142.
[10] Davydova, O. Medium. https://medium.com/@datamonsters/text-preprocessing-in-python-steps-tools-and-examples-bf025f872908, 2018, retrieved April 9, 2019.
[11] Islam, M., "Numeric Rating of Apps on Google Play Store by Sentiment Analysis on User Reviews", 1st International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), At Military Institute of Science and Technology, Dhaka, Bangladesh, 2014.
[12] Pamungkas, E. W. & Putri, D. G. P. "An Experimental Study of Lexicon-Based Sentiment Analysis on Bahasa Indonesia", Proceedings - 2016 6th International Annual Engineering Seminar, InAES 2016, pp. 28–31.
[13] Dhaoui, C., Webster, C.M., Tan, L.P., "Social Media Sentiment Analysis: Lexicon Versus Machine Learning", Journal of Consumer Marketing, Vol. 34 Issue: 6, 2017, pp.480-488
[14] Nurfalah, A. & Suryani, A. A. "*Analisis Sentimen Berbahasa Indonesia dengan Pendekatan Lexicon-Based Pada Media Sosial*", Jurnal Masyarakat Informatika Indonesia, 2(1), 2017, pp. 1–8.
[15] Suyanto. *Data Mining Untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika, 2019.
[16] Thelwall, M. SentiStrength. http://sentistrength.wlv.ac.uk/, 2012, retrieved April 2019.
[17] Xia, P. & Jiang, W. "Understanding the Evolution of Fine-Grained User Opinions in Product Reviews", Proceedings - 2018 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People and Smart City Innovations, SmartWorld/UIC/ATC/ ScalCom/CBDCo. IEEE, 2018, pp. 1335–1340.
[18] Katadata. "*Perilaku Konsumen E-Commerce*". Katadata Insight Center, November 2018, Indonesia.
[19] Han, SuHun. Googletrans: Free and Unlimited Google translate API for Python. 2018. https://py-googletrans.readthedocs.io/ en/latest/, retrieved Juli, 2019.

## BIBLIOGRAPHY OF AUTHORS

Erlina Halim, S.Kom., M.Kom., Graduate Informatics Engineering in STMIK Mikroskil, Medan

Dr. Ronsen Purba, M.Sc., Head of Department Informatics Engineering in STMIK Mikroskil, Medan

Andri, S.Kom., M.T.I., Head of Information System in STMIK Mikroskil, Medan