*Journal of Research in Artificial Intelligence for Systems and Applications (RAISA*
2025, Vol. 01, No. 01, E-ISSN: 3090-3157

Original Research

# Combining Rule Based, Lexicon Based and Support Vector Machine for Improve Accuracy in Sentiment Analysis of ChatGPT Usage

**Arham Zainul Abidin [1], Muhammad Ariful Furqon [2], Saiful Bukhori[3*]**

[1,2,3]Departement of Computer Science, University of Jember

## Abstract

ChatGPT is one of the Large Language Models (LLM) which is an artificial intelligence (AI) based chatbot. ChatGPT caused controversy in various news media due to its ability to interact and provide natural, human-like responses. This controversy created skepticism in society regarding the brand image. Therefore, a sentiment analysis was carried out specifically targeting Indonesian speaking Twitter users with a focus on the ChatGPT brand image. Various studies have been carried out to analyze user sentiment towards ChatGPT by analyzing tweets shared regarding ChatGPT using machine learning and deep learning. This research uses a combination of rule-based, Support Vector Machine (SVM), and lexicon-based methods. Rule-based classification uses emoticons and comparatives, while lexicon-based classification uses the BabelSenticNet lexicon. The data set, obtained through a Twitter crawl, initially consisted of 2,500 tweets, which were then cleaned to 1,728 tweets. After classification, model evaluation is carried out by comparing the performance between the proposed classification model and the constituent models. The proposed classification model has better performance, achieving 86.3% accuracy, 87.1% precision, 86.5% recall, and 86.6% f-measure. Sentiment predictions from 1,728 tweets resulted in 739 positive, 385 negative and 604 neutral sentiments. In conclusion, ChatGPT's brand image in Indonesia tends to be positive, even though there are differences in views and objective assessments.

## Keywords

ChatGPT, Large Language Models, sentiment analysis, lexicon based, rule based, Support Vector Machine.

## 1. Introduction

ChatGPT is one of the LLM (Kim et al. 2023). ChatGPT is an Artificial Intelligence based Chatbot developed by the OpenAI company (Roume liotis and Tselikas 2023). ChatGPT currently often causes controversy in various news media, including CNBC (Browne 2023). Because free text expressed in natural language often influences the prediction of action selection (Ben Porat et al. 2020). ChatGPT's controversy stems from its ability to interact and provide natural, human-like responses. The responses given are also accurate with the instructions written by the user. That is the reason why this chatbot is very popular and widely discussed on social media. ChatGPT can translate human language with high accuracy. Unlike conventional chatbots, ChatGPT is able to store conversation history to answer sub sequent questions, reject inappropriate requests, and refute false statements (Europol 2023). Some users also believe that ChatGPT has the potential to replace professions in the field of content writing, such as programmers, scriptwriters and journalists. Additionally, several companies have started integrating ChatGPT technology to optimize business potential. Based on the impressive performance of ChatGPT, there are some people who oppose the presence of ChatGPT, although the majority agree with its existence (Gondwe 2023). AI has the potential to enable agents to learn and perform tasks autonomously with superhuman

[*]Corresponding author: Saiful Bukhori

**Email addresses:**

arham.zaldin@gmail.com (Arham Zainul Abidin), ariful.furqon@unej.ac.id (Muh. Ariful Furqon), saiful.ilkom@unej.ac.id (Saiful Bukhori)

**Received:** 10 12 2024; **Accepted:** 24 12 2024; **Published:** 01 05 2025

performance (Retzlaff et al. 2024). The pros and cons related to the existence of ChatGPT give rise to public skepticism towards ChatGPT's brand image. Brand image is defined as a specific im- pression that is embedded in someone's memory regarding a business (Li et al. 2020). Brand image can be used as a variable to evaluate a product or other company attributes because it has a significant influence on an individual's intentions and decisions to carry out commercial activities.

Several studies have analyzed user sentiment towards ChatGPT by analyzing tweets shared related to ChatGPT using machine learning (Ouyang et al. 2024)( Nasayreh et al. 2024)(Sharma, Aggarwal, and Kumar 2023)(Sudheesh et al. 2023) and deep learning (M et al. 2023)(Diwali et al. 2023). Sentiment analysis tasks on aspects of conversation can provide de- tailed sentiment information that is useful for dialogue understanding and planning (Song et al. 2022). Several studies have also compared the use of machine learning and deep learning in analyzing sentiment of ChatGPT users (Baker and Utku 2023). This research has compared several deep learning and machine learning algorithms, including LSTM, Bi-LSTM, CNN, Gated Recurrent Unit, Random Forests and SVM. Experimental results show that the CNN-Bi-LSTM model trained by Fast Text outperforms other models in terms of accuracy.

This research will utilize rule-based, SVM, and lexicon-based approaches. The rule-based method employed consists of emoticon classification and comparative classification. According to the findings from (Xu, Yang, and Tian 2023), emoticons frequently appear in online forums due to their ease in expressing emotions. These emoticons can be emotional or comparative in nature. Emoticons and comparatives are identified by examining the lexical patterns in tweet using predefined rules. SVM is a supervised classification method that utilizes decision scores for data classification. These decision scores can be further processed using rule-based techniques to optimize classification. Meanwhile, the lexicon-based approach classifies sentiment based on the intensity of words using a lexicon. BabelSenticNet, which provides multi- lingual support, especially in Indonesian, is the lexicon used. BabelSenticNet offers thousands of sentiment-bearing words or concepts, organized using commonsense based algorithms built on Concept Net and additional algorithms based on Word- Net (Vilares et al. 2018).

Based on the outlined description, there is a need to analyze the sentiment on Indonesian language Twitter regarding the brand image of ChatGPT. This research is conducted to provide insights and perspectives on the brand image of ChatGPT to relevant stakeholders. By obtaining this information, stakeholders can use it as an indicator for decision-making, such as utilizing or improving ChatGPT. The research will combine rule-based, SVM, and lexicon-based methods to address the complexity of Twitter data structures and will only be combined if there is still data that does not meet the criteria. The rule-based method classifies using emoticons and comparatives because emoticons can clearly convey emotional references and comparatives can provide comparison information that can be interpreted with rules. Lexicon-based classification will use BabelSenticNet as it officially supports the Indonesian language and SVM will be employed for automatic classification due to its decision scores that can be processed using rule-based techniques.

## 2. Related Work

Sentiment analysis research on ChatGPT has been conducted by (Al-Khalifa et al. 2023), involving topic modeling and qualitative sentiment analysis of tweets using the identified topics. The data used in this research were sourced from the social media platform Twitter and were in English. The popularity of Twitter, also known as X nowadays, has led researchers and practitioners to frequently utilize its data to extract raw information from the public (Nazir and Wang 2023). As another ex- ample, Twitter data has been employed for testing sentiment classification methods (Li et al. 2023) and modeling public sentiment (Koonchanok, Pan, and Jang 2023). Therefore, this research will use Twitter as a data source to analyze the sentiment towards the brand image of ChatGPT.

Another sentiment analysis research on com- bined methods has been conducted by (Ashir 2021)(Braoudaki et al. 2020), successfully test- ing a combination of rule-based and lexicon-based sentiment analysis methods for Twitter data with complex text structures. Another research was carried out by (Cam et al. 2024) to enhance the quality of sentiment classification on Twitter using SVM and lexicon-based approaches. Subsequent research by (Adhiya et al. 2023) involved combin- ing lexicon-based and machine learning techniques for sentiment classification on Twitter data.

## 3. Research Method

The data crawling process will retrieve Indonesian language tweets with the keywords "chatgpt" and "chat gpt". To avoid subjectivity on the part of the researcher, data labeling will be conducted by experienced individuals, considering the aspects under investigation, namely brand image indica- tors, which include product quality, prestige, standards, services, and design. The analysis includes sentiment classification results analysis and tweet analysis. Therefore, specific characteristics will be processed during preprocessing to optimize classification and analysis outcomes. Preprocessing techniques such as Part of Speech (PoS) tagging will assign word classes to each token, allowing for sentiment-related parts to be identified. Tweet analysis will be expanded into topic analysis and the analysis of emoticon correlation with topics. The goal of topic analysis is to identify the sub- jects discussed in tweets, while the analysis of emoticon correlation with topics aims to under- stand the emotions related to a particular topic. Text pattern analysis will also be conducted to identify sentiment in tweets using lexical analysis to discover text patterns.

The classification process will combine three classification methods, namely rule-based, SVM, and lexicon-based. These methods will be combined if there is still data that remains unclassified in the previous methods. The sequence of combining these methods starts with rule-based, followed by SVM, and finally, lexicon-based. Rule-based is prioritized because it can effectively prioritize and process data containing emoticons and comparative key-words. The second sequence is SVM because it provides automatic classification but may not fully understand the context of emoticons. Meanwhile, the third sequence is lexicon-based, aiming to optimize the classification results of SVM. Therefore, SVM and lexicon-based must be in a specific order as shown in Figure 1.

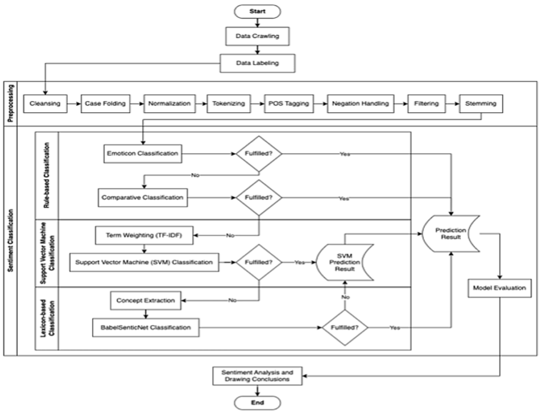


*Figure 1.* ***The Proposed Model***

## 3.1. Rule- based

In the initial stage of classification, data will be divided into training data and test data using 10-fold cross validation. The test data will be classified using rules and lexical patterns. Once the test data has been classified, it will be removed from the index to prevent reclassification, and the classification results will be stored in the prediction result variable. This classification will yield only positive and negative labels. The flow of the rule-based classification as shown in Figure 2.

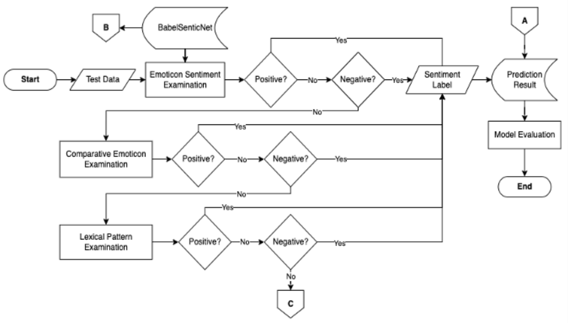


*Figure 2.* ***Rule-based Classification***

Emoticon classification, which involves classifying test data based on emoticons. BabelSenticNet will be utilized as it contains a list of emoticons. The rule is that if one or more positive emoticons are present and there are no negative emoticons then it is classified as positive. Conversely, if one or more negative emoticons are present and there are no positive emoticons, it is classified as neg- ative. If these conditions are not met, then the classification proceeds to the comparative classification. For example ”I like chatgpt 💗 😍” will be classified as positive, while ”chatgpt is good 👍 but a bit scary 😱” will not be classified based on emoticons.

Comparative classification, which involves classifying test data based on comparative key-words or comparative emoticons. The rule is that if a positive comparative emoticon is directly in front of the word “chatgpt” then it is classified as positive, and vice versa for negative. Meanwhile, if the pattern “chatgpt superior-of …” is found, it is classified as positive. On the other hand, if the pattern “… superior-from chatgp” is present, it is classified as negative. The same rule applies in reverse for “lose-to”. If these conditions are not met, the test data is then passed on to the SVM classification. For example, “google ❌ , chatgpt ✔ ” would be classified as positive. Another ex- ample related to the normalization of comparison symbols, such as “chatgpt < private school” be- coming “chatgpt lose-from private school” would be classified as negative.

## 3.2. Super Vector Machine-based

SVM will classify the remaining test data in the index and store the results in the SVM pre-diction result variable. Once the method combination is completed, the SVM prediction result variable will be passed to the main prediction result variable. The SVM classification produces positive, negative, and neutral labels. The flow of the SVM classification as shown in Figure 3.

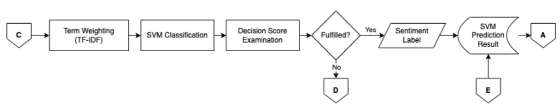


*Figure 3.* ***SVM Classification***

Term weighting (TF-IDF), which involves weighting term using TF-IDF. The resulting numerical data from the weighting process will be used in the SVM classification. SVM classification, which involves classifying test data using the SVM method. The examination of decision scores in the classification results is performed using the decision function(). The threshold for the decision score is determined by inspecting the dot plot graph of decision scores and selecting the lowest value from the majority data set. The classification results with decision scores meeting the threshold are stored in the SVM prediction variable. If the thresh-old is not met, tweet is then passed on to the lexicon-based classification.

## 3.3. Lexicon- based

This classification will use textual test data with the same index as the forwarded numerical test data. The lexicon-based classification will produce only positive and negative labels. The flow of the lexicon-based classification as shown in Figure 4.

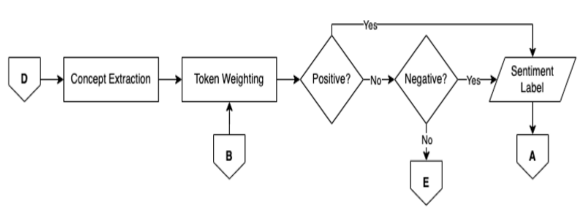


*Figure 4.* ***Lexicon-based Classification***

Concept extraction, which involves extracting concepts in the form of tokens from a document. This is done because the BabelSenticNet lexicon only weights tokens individually. BabelSentic- Net classification, which involves weighting tokens with sentiment scores according to the BabelSen- ticNet lexicon. The score calculation is done using the average. The threshold to be used is 0.33 and -0.33, which are obtained from the lowest sentiment weights in the BabelSenticNet lexicon. Some studies use thresholds close to these values, such as in the research (Nasayreh et al. 2024), which also classifies with a lexicon. The criteria are, if the average is 0.33, it is classified as positive, whereas if it is -0.33, it is classified as negative. The classi- fication results meeting these criteria will replace the SVM results. Otherwise, the previous SVM results will be retained. For example, in the tweet "chatgpt more smart" there is a word composition of "chatgpt"=0, "more" = 0, and "smart" = 1. So, the calculation is (0+0+1)/3=0.33 and it is classified as positive.

### 3.4. Modelling

This stage will analyze the model's performance by calculating the averages of accuracy, precision, recall, and f-measure, both before and after the combination of classification methods. Therefore, a performance comparison will be conducted be- tween the proposed classification models, namely rule-based, SVM, and lexicon-based with their constituent classification models, namely, rule-based single classification and combined classification of rule-based with SVM.

### 3.5. Sentiment Analysis and Results

This stage will analyze the presented sentiment classification results and examine tweets to delve deeper into the meaning expressed by the opinion holders. The entire analysis process will be con- ducted manually. The classification results analysis will summarize which sentiment is predominant, followed by other sentiments. Meanwhile, the tweet analysis will involve taking a sample from the dataset. Each stage of tweet analysis will consider brand image indicators. The stages of tweet analysis include: (1) The sentiment of sample tweet will be manually classified and validated by identifying words or emoticons conveying sentiment with the assistance of POS tagging, (2) The topics in the sample tweet will be identified to measure customer satisfaction and identify issues related to the product, (3) If there are emoticons, those emoticons will also be analyzed to identify emotions related to a topic.

## 4. Results and Analysis

### 4.1. Data Crawling

The data crawling stage on Twitter is performed using the Twitter API. The latest rules of the Twitter API limit users to fetching data once per day. Therefore, data crawling is conducted gradually. The search queries entered are "chatgpt" and "chat gpt". The filters used during the search are "lang:id" and "-filter:retweets", meaning the search results include Indonesian-language tweets and exclude retweets. The crawling is done incrementally from 10 April 2023 to 10 May 2023. The total data obtained from this data crawling stage is 2500 tweets. Table 1 presents some tweets obtained from the data crawling stage.

***Table 1. Data Crawling***

| *No* | *Tweets* |
|---|---|
| 1 | @idextratime Enak banget yah sekarang udah ada chatgpt |
| 2 | @collegemenfess chatgpt adalah kunci 👍 |
| 3 | ngga susah kan ada chatgpt 😃 https://t.co/ZCTuzGJ1jZt |
| 4 | @ReplyGPT @KenWalady @txtdari sisange HAAOWKAWAOKW, BAHKAN AI CHATGPT AJA NGAK BISA JAWAB |
| 5 | dosen bahas chatgpt dll, aku (((bombas- tic criminal side eye))) |

### 4.2. Data Labeling

The labeling is done by considering indicators of the brand image as the evaluated aspects, namely product quality, prestige, standards, services, and design. The labels used are positive, negative, and neutral. Before labeling, data will be formatted for readability by labeler and validator. Duplicate data, empty data, and data in foreign languages are also cleaned to minimize labeling costs. A total of 1798 tweets were labeled.

### 4.3. Model Development

#### 4.3.1. Data Preprocessing

The data preprocessing stage is performed similarly to the general classification modeling. In this research, there are several additional preprocessing stages to handle the complexity of Twitter data in informal text form, namely normalization, POS tagging, and negation handling. So, the preprocessing stages include cleansing, case folding, normalization, tokenizing, POS tagging, negation handling, filtering, and stemming. The preprocessing stages begin by cleaning the data from usernames, emails, links, white spaces, specific punctuation, and certain characters. Some characters are reformatted to identify emoticons and comparative punctuation. Removal and formatting of characters are done using regex and several functions. The data is also formatted to lowercase for consistency. Subsequently, data selection is carried out for strings containing "gpt" and non-duplicate, resulting in a total of 1728 tweets. Table 2 presents the number of tweets for each sentiment in this stage.

***Table 2. Number of tweets on each sentiment***

| Label | Total |
|---|---|
| Positive | 681 |
| Negative | 400 |
| Neutral | 647 |

For other preprocessing stages, such as stemming, it will be performed as usual in modeling and adapted to handle emoticons. Specific preprocess- ing stages, such as normalization, POS tagging, negation handling, and filtering, will be explained below. All preprocessing stages are conducted sequentially:

1. Normalization; The normalization stage begins by creating a dictionary of normalized words, which are more appropriate and standard. The creation of the word dictionary is done manually and automatically with the assistance of the KBBI API. The word dictionary includes original terms, foreign terms, slang words, repeated words, synonyms, and non-alphabetic characters. Some non-alphabetic characters are comparative symbols. So, the symbol “<” becomes the keyword “kalah-dari” and “>” becomes “unggul-dari”. After the word dictionary is created, the data is matched with the dictionary and words matching the dictionary will be replaced. Table 3 presents the results of normalization.

***Table 3. POS Tagging Process***

| *Before* | *After* |
|---|---|
| Tidak susah kan ada chatgpt 😃 | (’Tidak’, ’NEG’), (’susah’, ’JJ’), (’kan’, ’RP’), (’ada’, ’VB’), (’chatgpt’, ’FW’), (’ 😃’, ’SYM’) |

2. POS Tagging: The process of assigning word classes to each token is carried out using Kumparan’s NLP Services library (nlp.id). This library performs both POS tagging and tokenizing processes. The results of POS tagging can be seen in Table 4. In Table 4 “After” column, it can be observed that each token has its respective word class. The word class “NEG” indicates negation, “JJ” stands for adjectives, “RP” represents particles, “VB” denotes verbs, “FW” signifies foreign terms, and “SYM” indicates symbols.

***Table 4. Negation handling process***

| *Before* | *After* |
|---|---|
| [(’tidak’, ’NEG’), (’susah’,JJ’), (’kan’, ’RP’), (’ada’, ’VB’), (’chat- gpt’, ’FW’), (’😃’, ’SYM’)] | [(’tidak’,’NEG’), (’susah’,’JJ’, ’NEG’),(’kan’, ’RP’), (’ada’, ’VB’), (’chatgpt’, ’FW’), (’ 😃’, ’SYM’) |

3. Negation Handling: To preserve the original sentiment, it is necessary to add special codes to words in front of negation words because the classification model reads tokens individually, allowing the model to differentiate between the word “difficult” in [difficult] and the word “difficult” in [not difficult]. The code involves adding the word class “NEG” to the word in front of it. The results can be seen in Table 5.
4. Filtering: The removal of irrelevant tokens is necessary to lighten the classification process. The removed tokens are those contained in the NLTK library and the additional dictionary. The additional dictionary is created manually to suit the existing data. Negation words are also removed at this stage because their word classes have been added to the words in front of them, making them unnecessary. Meanwhile, words with the word classes “ADV”, “JJ” and “VB” will not be deleted. The results of this stage can be seen in Table 6.

***Table 5. Negation handling process***

| *Before* | *After* |
|---|---|
| [(’tidak’, ’NEG’), (’susah’,JJ’), (’kan’, ’RP’), (’ada’, ’VB’), (’chat- gpt’, ’FW’), (’😃’, ’SYM’)] | [('tidak', 'NEG'), ('susah', 'JJ', 'NEG'), ('kan', 'RP'), ('ada', 'VB'), ('chatgpt', 'FW'), ('😌', 'SYM')] |

***Table 6. Filtering process***

| *Before* | *After* |
|---|---|
| [(’tidak’, ’NEG’), (’susah’,JJ’), (’kan’, ’RP’), (’ada’, ’VB’), (’chat- gpt’, ’FW’), (’😃’, ’SYM’)] | [('susah', 'JJ', 'NEG'), ('ada', 'VB'), ('chatgpt', 'FW'), ('😌', 'SYM')] |

### 4.3.2. Sentiment Classification

The classification stage will be processed using 10-fold cross-validation. Each fold will undergo classification, starting with rule-based. If there are still unclassified tweet, Support Vector Machine classification will be performed. However, if there is tweet with a low Support Vector Machine decision score, it will be passed to lexicon-based. The results of lexicon-based classification will replace the labels from the Support Vector Machine classification results if valid labels are obtained. If not, the labels from the previous Support Vector Machine classification results will be retained.

- Rule-based Classification: In the rule-based classification, there are two classification processes, namely classification using emoticons and comparative classification. The classification rules using emoticons are shown in Table 7.

*Table 7. Emoticons classification rules*

| No | Classification |
|---|---|
| 1 | If it contains one or more positive labels and no negative labels, tweet will be classified as positive. |
| 2 | If it contains one or more negative labels and no positive labels, tweet will be classified as negative. |

If any rule in Table 7 is satisfied, tweet will store the label from the classification results using emoticons. However, if the rules are not satisfied, tweet will be classified comparatively. The comparative classification rules are presented in **Error! Reference source not found.**.
If any rule in Table 8 is satisfied, tweet will store the label from the comparative classification results. However, if the rules are not satisfied, tweet will be forwarded to the SVM classification.

*Table 8. Comparative classification rules*

| No | Classification |
|---|---|
| 1 | If the string "chatgpt" is behind the emoticon ❌ , tweet will be classified as positive. |
| 2 | If the string "chatgpt" is behind the keword "unggul dari", tweet will be classified as positive. |
| 3 | If the string "chatgpt" is in front of the keyword "unggul dari", tweet will be clas- sified as negative. |
| 4 | If the string "chatgpt" is behind the emoticon ✔, tweet will be classified as negative. |
| 5 | If the string "chatgpt" is behind the key- word "kalah dari", tweet will be classified as negative. |
| 6 | I If the string "chatgpt" is in front of the keyword "kaah dari", tweet will be clas- sified as positive. |

- SVM Classification: Tweet forwarded from the rule-based classification will undergo the process of term weighting using TF-IDF. Af- ter the term are weighted, classification with SVM will be conducted. The obtained pre- diction results in this classification will be examined against the decision score. The plot of the decision score in SVM classification as shown in Figure 5.
-

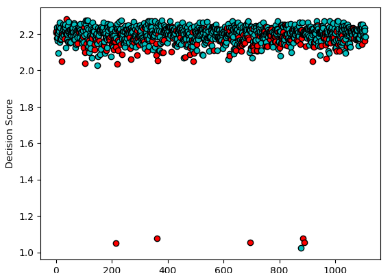


*Figure 5. Dot plot SVM decision score*

In Figure 5, it can be observed that the red dots, indicating misclassifications, tend to be below the score of 2.1. Therefore, a rule will be implemented for the SVM classification, stating that tweet with a decision score <2.1 will undergo reclassification using lexicon-based.

- Lexicon-based Classification; Tweet forwarded from the SVM classification will be formatted into a string for concept extraction before lexicon-based classification. Data formatting will be done by extracting the tweet before the TF-IDF process whose indices match those of the tweet forwarded from the Support Vector Machine classification. Once the textual tweet is obtained, tweet will be tokenized. Each token will have its sentiment score identified using BabelSen- ticNet, and then these scores will be summed to obtain the total sentiment score. After obtaining the total sentiment score, an average will be calculated for this total sentiment score. Subsequently, the average score will be applied to the rules in Table 9. If any rule in Table 9 is satisfied, the label from the lexicon-based classification results will replace the label from the SVM classification results. However, if the rules are not satisfied, tweet will retain the label from the previous SVM classification results.

*Table 9. Emoticons classification rules*

| No | Classification |
|---|---|
| 1 | If the average score is $\geq 0.33$ then tweet will be classified as positive. |
| 2 | If the average score is $\leq -0.33$ then tweet will be classified as negative. |

After the classification stage is complete, the aggregation of all classification results is per-formed. The aggregated classification results as shown in Figure 6


*Figure 6. Classification results*

## 4.4. Model Evaluation

From the constructed classification model, a performance comparison will be conducted with the forming classification models, namely rule-based single classification and combined classification of rule-based with Support Vector Machine. The comparison will be presented using evaluation metrics, taking the average classification performance from the entire 10-fold execution. The first forming classification model is rule-based single classification. The performance of this classification model is presented in Table 10.

The second forming classification model is the combination of rule-based and Support Vector Machine. The performance of the combined rule-based and Support Vector Machine classification model is presented in Table 11.

***Table 10. Performance of rule-based classification***

| Metrics | Average |
|---|---|
| Accuracy | 0.722 |
| Precision | 0.848 |
| Recall | 0.714 |
| F-measure | 0.725 |

***Table 11. Performance of combined rule-based and SVM***

| Metrics | Average |
|---|---|
| Accuracy | 0.86 |
| Precision | 0.869 |
| Recall | 0.862 |
| F-measure | 0.864 |

Next is the proposed combined classification model in this research, namely the combination of rule-based, Support Vector Machine, and lexicon-based classification. The performance of this classification model can be seen in Table 12

***Table 12. Performance of combined rule-based, SVM. and lexicon-based***

| Metrics | Average |
|---|---|
| Accuracy | 0.863 |
| Precision | 0.871 |
| Recall | 0.865 |
| F-measure | 0.866 |

Based on the presented performance of classification models, it is found that the performance of the combined rule-based, Support Vector Machine, and lexicon-based classification model is superior to its forming models. The proposed classification model excels in accuracy with an average of 86.3%, average precision of 87.1%, average recall of 86.5%, and average f-measure of 86.6%.

## 4.5. Sentiment Analysis and Drawing Conclusions

After successfully building the classification model and obtaining the classification results, the next step is to analyze and discuss the results from the previous stages. The analysis is done manually by reading the presented tables.

### 4.5.1. Classification Results Analysis

Based on the classification results presented in **Error! Reference source not found.**, it is known that the sentiment towards the ChatGPT brand image in Indonesia on the Twitter platform is generally positive. ChatGPT receives positive sentiment of 42.8% from 1728 data, the highest among other sentiments. As a new technology, ChatGPT is not without its share of both positive and negative sentiments. This is evident in the emergence of negative sentiment at 22.3%. Meanwhile, others maintain a neutral stance and assess objectively, as reflected in the occurrence of neutral sentiment at 34.9%.

### 4.5.2. Tweet Analysis

In the tweet analysis stage, a sample will be taken from the dataset. POS tagging will be used in this stage to assist the analysis process. Before that, the sample will be normalized first to obtain optimal results. BabelSenticNet lexicon will also be used for emoticon analysis. The sample to be analyzed can be seen in Table 13.

***Table 13. Performance of combined rule-based, SVM. and lexicon-based***

| Original Tweet | POS Tagging | Prediction |
|---|---|---|
| chatGPT ini lebih powerful dr brainly 😌 | chatgpt_FW ini_DT lebih_ADV kuat_JJ dari_IN brainly_NN 😌_SYM | Positive |

In Table 13, a tweet classified as positive has been presented. Based on the POS tagging results of this tweet, there is the word “kuat” as a positively sentimental word, which is also marked as an adjective (JJ), followed by the emoticon “😌” indicating satisfaction. Therefore, this tweet is correctly classified as positive. Additionally, the user expresses satisfaction with the quality of ChatGPT compared to its competitors, which is reinforced by the presence of the emoticon “😌” indicating satisfaction with the quality of ChatGPT.

# 5. Conclusion

Based on the conducted research, the following conclusions can be drawn. Implementation of the sentiment classification model by combining rule-based, Support Vector Machine, and lexicon-based methods can be effectively performed. This is evident from the superior performance of the proposed classification model compared to its constituent models, namely rule-based single classification and combined classification of rule-based with Support Vector Machine. The proposed classification model, which combines rule-based, Support Vector Machine, and lexicon-based, achieved an accuracy score of 86.3%, precision of 87.1%, recall of 86.5%, and f-measure of 86.6%.

Sentiment analysis of Twitter users towards the ChatGPT brand image using rule-based, Support Vector Machine, and lexicon-based on 1728 data resulted in 739 data or 42.8% being classified as positive, 385 data or 22.3% as negative, and 604 data or 34.9% as neutral. From these results, it can be concluded that the ChatGPT brand image in Indonesia tends to be positive, although some others express opposing views and assess it objectively.

## References

[1] J. K. Kim, M. Chua, M. Rickard, and A. Lorenzo, "ChatGPT and large language model (LLM) chatbots: The current state of acceptability and a proposal for guidelines on utilization in academic medicine," J. Pediatr. Urol., vol. 19, no. 5, p. 607, 2023, doi: 10.1016/j.jpurol.2023.07.007.

[2] K. I. Roumeliotis and N. D. Tselikas, "ChatGPT and Open-AI Models: A Preliminary Review," Futur. Internet, vol. 15, no. 6, 2023, doi: 10.3390/fi15060192.

[3] R. Browne, "All you need to know about ChatGPT, the A.I. chatbot that's got the world talking and tech giants clashing.," CNBC, pp. 1–9, 2023. [Online]. Available: https://www.cnbcafrica.com/2023/all-you-need-to-know-about-chatgpt-the-a-i-chatbot-thats-got-the-world-talking-and-tech-giants-clashing/

[4] Europol, "ChatGPT - The impact of Large Language Models on Law Enforcement, a Tech Watch Flash Report from the Europol Innovation Lab," Eur. Public Inf., 2023, [Online]. Available: https://www.europol.europa.eu/cms/sites/default/files/documents/Tech Watch Flash - The Impact of Large Language Models on Law Enforcement.pdf

[5] G. Gondwe, "CHATGPT and the Global South: how are journalists in sub-Saharan Africa engaging with generative AI?," Online Media Glob. Commun., vol. 2, no. 2, pp. 228–249, 2023, doi: 10.1515/omgc-2023-0023.

[6] Y. Li, W. Teng, T. T. Liao, and T. M. Y. Lin, "Exploration of patriotic brand image: its antecedents and impacts on purchase intentions," Asia Pacific J. Mark. Logist., vol. 33, no. 6, pp. 1455–1481, 2020, doi: 10.1108/APJML-11-2019-0660.

[7] T. Ouyang, A. MaungMaung, K. Konishi, Y. Seo, and I. Echizen, "Stability Analysis of ChatGPT-based Sentiment Analysis in AI Quality Assurance," pp. 1–8, 2024, [Online]. Available: http://arxiv.org/abs/2401.07441

[8] A. Nasayreh, R. Emhamed, A. L. Mamlook, G. Samara, and D. A. Bi, "Arabic Sentiment Analysis for ChatGPT Using Machine Learning Classification Algorithms : A Hyperparameter Optimization Technique," ACM Trans. Asian Low-Resource Lang. Inf. Process., 2024, doi: 10.1145/3638285.

[9] S. Sharma, R. Aggarwal, and M. Kumar, "Mining Twitter for Insights into ChatGPT Sentiment: A Machine Learning Approach," 2nd IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. ICDCECE 2023, no. June, pp. 2–5, 2023, doi: 10.1109/ICDCECE57866.2023.10150620.

[10] R. Sudheesh et al., "Analyzing Sentiments Regarding ChatGPT Using Novel BERT: A Machine Learning Approach," Inf., vol. 14, no. 9, pp. 1–29, 2023, doi: 10.3390/info14090474.

[11] A. K. M, N. Nandhini, G. Kavitha, and N. Ezra, "ChatGPT in Future Data Analytics," Eur. Chem. Bull., vol. 12, no. 8, pp. 3423–3433, 2023.

[12] A. Diwali, K. Saeedi, K. Dashtipour, M. Gogate, E. Cambria, and A. Hussain, "Sentiment Analysis Meets Explainable Artificial Intelligence: A Survey on Explainable Sentiment Analysis," IEEE Trans. Affect. Comput., pp. 1–11, 2023, doi: 10.1109/TAFFC.2023.3296373.

[13] M. R. Baker and A. Utku, "Unraveling user perceptions and biases: A comparative study of ML and DL models for exploring twitter sentiments towards ChatGPT," J. Eng. Res., no. November, 2023, doi: 10.1016/j.jer.2023.11.023.

[14] L. Xu, X. Yang, and S. Tian, "A Study on the Role of Internet Emoticons in Business Communication from the Perspective of Symbolic Interactionism," Athens J. Mass Media Commun., vol. 9, no. 3, pp. 161–184, 2023, doi: 10.30958/ajmmc.9-3-2.

[15] D. Vilares, H. Peng, R. Satapathy, and E. Cambria, "BabelSenticNet: A Commonsense Reasoning Framework for Multilingual Sentiment Analysis," Proc. 2018 IEEE Symp. Ser. Comput. Intell. SSCI 2018, pp. 1292–1298, 2018, doi: 10.1109/SSCI.2018.8628718.

[16] S. Al-Khalifa, F. Alhumaidhi, H. Alotaibi, and H. S. Al-Khalifa, "ChatGPT across Arabic Twitter: A Study of Topics, Sentiments, and Sarcasm," Data, vol. 8, no. 11, 2023, doi: 10.3390/data8110171.

[17] A. Nazir and Z. Wang, "A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges," Meta-Radiology, vol. 1, no. 2, p. 100022, 2023, doi: 10.1016/j.metrad.2023.100022.

[18] X. Li et al., "Are ChatGPT and GPT-4 General-Purpose

Solvers for Financial Text Analytics? A Study on Several Typical Tasks," Proc. 2023 Conf. Empir. Methods Nat. Lang. Process., pp. 408–422, 2023, doi: 10.18653/v1/2023.emnlp-industry.39.

[19] R. Koonchanok, Y. Pan, and H. Jang, "Tracking public attitudes toward ChatGPT on Twitter using sentiment analysis and topic modeling," pp. 1–21, 2023, [Online]. Available: http://arxiv.org/abs/2306.12951

[20] A. M. Ashir, "A Generalized Method for Sentiment Analysis across Different Sources," Appl. Comput. Intell. Soft Comput., vol. 2021, 2021, doi: 10.1155/2021/2529984.

[21] A. Braoudaki, E. Kanellou, C. Kozanitis, and P. Fatourou, "Hybrid Data Driven and Rule Based Sentiment Analysis on Greek Text," Procedia Comput. Sci., vol. 178, no. 2019, pp. 234–243, 2020, doi: 10.1016/j.procs.2020.11.025.

[22] H. Cam, A. V. Cam, U. Demirel, and S. Ahmed, "Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers," Heliyon, vol. 10, no. 1, p. e23784, 2024, doi: 10.1016/j.heliyon.2023.e23784.

[23] S. Adhiya, S. Shirbhate, A. Kadu, A. Kalambe, and T. Mohod, "Combining Lexicon based and Machine Learning based Methods for Twitter Sentiment Analysis," Int. Res. J. Eng. Technol., pp. 191–195, 2023.