West Science Social and Humanities Studies
Vol. 02, No. 07, July 2024, pp. 1237-1245

Enhancing Customer Experience in E-commerce through Lexicon and TextBlob Sentiment Analysis

Hartatik
Department of Informatics Engineering, Vocational School, Sebelas Maret University, Surakarta, Indonesia

Article Info

Article history:

Received July, 2024
Revised July, 2024
Accepted July, 2024

Keywords:

Text Mining
Sentiment
Analysis Lexicon
TextBlob
Data
Science

ABSTRACT

This study evaluates customer satisfaction in business and e-commerce using sentiment analysis based on Indonesian Lexicon and TextBlob. The method used in this study is an explorative quantitative approach with sentiment analysis techniques that compare the Lexicon and TextBlob methods in processing customer review data. The analysis results show the dominance of the neutral sentiment category, with Lexicon producing around 1400 neutral reviews, 1000 positive reviews, and less than 200 negative reviews, while TextBlob shows more than 2000 neutral reviews with less than 500 positive reviews and almost no negative reviews. These findings reveal that the Lexicon method is more sensitive in detecting positive sentiment than TextBlob, which tends to be conservative. The implication of this study is the importance of choosing the right sentiment analysis method to improve customer service strategies. With an accuracy score of 78.52%, precision of 68.11%, and F1-Score of 63.54%, this analysis provides practical insights into how companies can effectively interpret customer sentiment to improve service quality and overall customer satisfaction.

Corresponding Author:

Name: Hartatik, S.Si, M.Si, M.Kom
Institution: Department of Informatics Engineering, Vocational School, Sebelas Maret University, Surakarta, Indonesia.
Email: hartatik.uns@gmail.com

1. INTRODUCTION

In recent years, the rapid advancement of information technology and artificial intelligence has brought about a significant transformation in business operations, particularly in the realm of customer service in the business and e-commerce industries. One technology that has emerged as a game-changer is machine learning-based sentiment analysis, a branch of Natural Language Processing (NLP) [1], [2]. This powerful tool enables companies to continuously process and understand customer opinions, emotions, and attitudes extracted from textual data, including product reviews, social media interactions, and customer feedback [3]. The potential of this technology to revolutionize customer service is truly inspiring.

Sentiment analysis powered by machine learning can process vast amounts of textual data in real time, identifying trends and patterns that traditional methods might overlook [4]. With these insights, businesses can proactively address customer needs and complaints, significantly improving customer satisfaction and loyalty, which drives sustainable revenue growth. Companies effectively using sentiment analysis can thus achieve greater adaptability, respond more efficiently to market changes, and sustain continuous growth in a highly competitive environment [5] Designing a sales prediction model in tourism industry and hotel recommendation based on hybrid recommendation.

Previous research underscores the efficacy of sentiment analysis in enhancing business performance across various sectors, particularly highlighting its capability to detect consumer sentiment and predict market trends quickly. [4]demonstrated how sentiment analysis contributes significantly to customer loyalty and revenue growth by enabling timely and informed decision-making [6]. Additionally, studies by [7] illustrated that automated sentiment analysis outperforms traditional manual methods by improving accuracy and reducing processing time, thus enabling proactive and effective customer relationship management [8].

Building on these prior findings, this research delves deeper into the practical implications of sentiment analysis in business and e-commerce. It specifically focuses on sentiment computation using TextBlob and Lexicon-based approaches, chosen for their proven efficiency and adaptability in accurately interpreting customer feedback [9], [10]. By evaluating and comparing the performance of these two methods, this study provides actionable insights into how businesses can effectively leverage sentiment analysis technologies to refine service quality and strategic decision-making processes. This emphasis on practical applications aims to equip the audience with a deeper understanding of the potential of sentiment analysis in their respective fields [11], [12].

Despite its demonstrated advantages, sentiment analysis presents a set of challenges, including data privacy concerns, algorithmic biases, linguistic complexity, and cultural differences. It is crucial for organizations to comprehensively manage these issues to fully realize the technology's potential in an ethical and effective manner. By addressing these challenges head-on, businesses can maximize the positive impacts of sentiment analysis on customer satisfaction and long-term competitiveness.

The research focuses on the critical role of machine learning-based sentiment analysis in enhancing service quality and strategic responsiveness in business and e-commerce. By effectively integrating advanced analytical tools like TextBlob and Lexicon-based methods into their strategic operations, organizations can significantly improve customer satisfaction and achieve sustainable competitive advantages in rapidly evolving digital marketplaces [13], [14].

The contribution of this study is the selection of sentiment analysis methods the use of these two methods lies in the needs of the business and e-commerce industry for a quick response to market dynamics, where customer sentiment changes rapidly and needs to be analyzed accurately and efficiently. The selection of these two sentiment analysis methods is a novelty in research, considering that previous studies generally only use one approach, so this study is able to offer a new perspective in the use of sentiment analysis technology.

The remainder of this paper is structured as follows. Section 2 reviews relevant literature on the concept of service quality and the application of NLP techniques in this domain, namely service analysis. Section 3 describes the methodology used to implement the proposed rumor detection system, including data collection, pre-processing, and the machine learning models employed. Section 4 presents the results of the experiments, followed by a discussion of the findings in Section 5. Finally, Section 6 concludes the paper and outlines directions for future research.

2. LITERATURE REVIEW

2.1 Concept Quality of Service

Quality of Service (QoS) is an important aspect in the business and e- commerce industry because it directly impacts customer satisfaction, loyalty, and overall business success. Parasuraman et al. (1985) define service quality as the gap between customer expectations and their perceptions of actual service performance. This concept emphasizes the importance of consistently delivering superior service that meets customer expectations [15]. In addition, QoS is usually measured through several key dimensions, such as reliability, responsiveness, assurance, empathy, and physical elements, collectively shaping the overall customer experience [16].

In the context of digital business and e-commerce, the concept of service quality has undergone significant changes along with the development of information technology. The digitization of services has increased customer expectations for fast response times, personalized experiences, and seamless interactions [17]. Therefore, businesses need to leverage data analytics and artificial intelligence to understand customer needs more deeply and dynamically adjust service offerings to improve customer satisfaction.

2.2 *Sentiment Analysis*

Sentiment analysis, or opinion mining, is the computational study of opinions, sentiments, emotions, and attitudes expressed in textual data (Liu, 2012). The primary goal of sentiment analysis is to identify and classify the polarity of text data into positive, negative, or neutral categories. The increasing amount of user-generated content, especially from social media platforms and customer reviews, makes sentiment analysis a valuable tool for businesses seeking to understand customer preferences and attitudes [18].

Various sentiment analysis methods

have been developed, including machine learning, lexicon-based, and hybrid approaches that combine the two. Machine learning techniques such as supervised learning algorithms are widely used because of their accuracy and adaptability in classifying sentiment [18]. On the other hand, lexicon-based methods are in demand because of their transparency and ease of implementation, using a dictionary of words that have been labeled with sentiment to determine sentiment scores [15], [19].

2.3 *Sentiment Analysis with Textblob and Lexicon*

Sentiment analysis has emerged as an essential analytical method within the domain of natural language processing (NLP), particularly to understand consumer emotions, opinions, and perceptions towards products and services. Among various sentiment analysis methods, TextBlob and lexicon-based approaches stand out for their accessibility, interpretability, and efficacy [20] Anomaly detection through enhanced sentiment analysis on. TextBlob, a Python library, integrates multiple NLP functions, including sentiment analysis, leveraging machine learning algorithms alongside heuristic rules. It classifies textual data into positive, neutral, or negative sentiments by evaluating linguistic patterns and contextual meanings within sentences [21]. The TextBlob approach benefits accuracy, particularly useful in scenarios requiring quick and reliable insights into customer opinions [22].

The sentiment polarity score in TextBlob is computed using the following formula in (1).

$$\{n\}$$
$$Senti_{\{polarity\}} = \{\Sigma\ polarity(word_i)\}\{n\} \qquad (1)$$

Where:

1. $Senti_{\{polarity\}}$ is the polarity score of the individual word.
2. n is the total number of words in the analyzed text.

Conversely, lexicon-based sentiment analysis utilizes pre-compiled dictionaries or lexicons containing words explicitly annotated with sentiment polarity scores. The sentiment of a text is computed by aggregating the sentiment values of individual words found within the lexicon. This method is appreciated for its transparency and ease of customization, making it particularly advantageous in domain-specific analyses where specialized vocabulary significantly influences sentiment interpretation.

Lexicon-based sentiment analysis typically follows this general formula in equation (2).

$$Senti_{\{score\}} = \sum_{\{i=1\}}^{\{n\}} score(word_i) \qquad (2)$$

Where:

1. $Senti$ is the lexicon-assigned sentiment score for the word .
2. $word_i$ is the total number of words identified in the lexicon from the analyzed text.

Previous studies highlight both methods' strengths and limitations. TextBlob has been noted for efficiently capturing contextual nuances yet can be conservative, leading to a higher prevalence of neutral classifications. On the other hand, lexicon- based approaches are recognized for their sensitivity in detecting explicit positive and negative sentiments but may face challenges in interpreting contextual or nuanced language.

Lexicon-based methods match words in a text with a sentiment dictionary containing scores for each word or phrase. The sentiment polarity of a text is then determined by summing the scores of the identified words. According to Machova et al. (2020), lexicon-based approaches have advantages in terms of interpretability, transparency, and ease of updating the dictionary for a particular context or domain

[23]. However, this method also has limitations in handling texts containing sarcasm, context-dependent meanings, or field-specific jargon [24]. Therefore, further development and use of hybrid approaches with machine learning methods are often needed to improve the accuracy of sentiment analysis results.

3. METHODS

This study uses an exploratory design that is appropriate to explore the effectiveness of machine learning-based sentiment analysis in the context of the business and e-commerce industry. This design was chosen because of its ability to deeply explore the potential of sentiment analysis in improving service quality by interpreting customer textual data. The study focuses on processing text data, especially customer reviews, social media comments, and feedback related to products or services provided by customers.

The data used in this study is kaggle.com data. The data obtained includes user, rating, and text review data. Based on the dataset used, this study uses two sentiment analysis methods: TextBlob and Lexicon-based. These two methods are selected based on the need to produce accurate and comprehensive sentiment analysis and as a form of research novelty. The use of these two methods is expected to provide a clear comparison of the effectiveness of each technique in producing accurate sentiment classification and strengthening the validity of research findings.

Sentiment analysis using the TextBlob method is carried out with the TextBlob library in Python which integrates machine learning algorithms and heuristic approaches. The advantage of TextBlob lies in its ability to automatically identify text sentiment based on language patterns and sentence context, so that it can provide fast and relevant analysis results. The analysis process includes data preprocessing, sentiment classification, and identification of sentiment patterns and trends to generate strategic insights for business in Figure 1.

```python
def get_sentiment(text):
    analysis = TextBlob(text)
    score = analysis.sentiment.polarity
    if score > 0:
        return "Positif"
    elif score < 0:
        return "Negatif"
    else:
        return "Netral"
```

Figure 1. Script of Sentiment Analysis

As a comparative approach, this study also uses the Lexicon-based method, which works by using a special sentiment dictionary containing words that have a certain sentiment score. This method provides transparent and systematic sentiment analysis by matching words in the text with the available dictionary [25]. The combination of TextBlob and Lexicon-based methods is used to improve the accuracy of sentiment analysis results and provide a deeper understanding of the nuances of customer emotions in their reviews and feedback.

4. RESULTS AND DISCUSSION

4.1 Sentiment Distribution Using Indonesian Lexicon

Based on sentiment analysis conducted using the Indonesian Lexicon method, it was found that most customer reviews were in the neutral sentiment category, with the number of reviews reaching around 1400 reviews. Positive sentiment was in second place with around 1000 reviews, while negative sentiment was relatively small, under 200 reviews. This distribution shows a general tendency for customers who do not explicitly show satisfaction or dissatisfaction in their reviews, but rather tend to be neutral. This pattern indicates that the service or product has met customer expectations but is not strong enough to produce an explicit positive impression in Table 1. Lexicon-based sentiment analysis using K-Nearest Neighbors (KNN), which is generalizing and domain independence, allows sentiment detection.

Table 1. Labelling of sentiment analysis

Label	Rule
Positive	Weight > 0
Neutral	Weight = 0
Negative	Weight < 0

Below in Figure 2 is a wordcloud visualization for analysis using lexicon.

Figure 2. Word CLoud of Sentiment Analysis Using Lexicon (a) Positive (b) Neutral (c) Negative

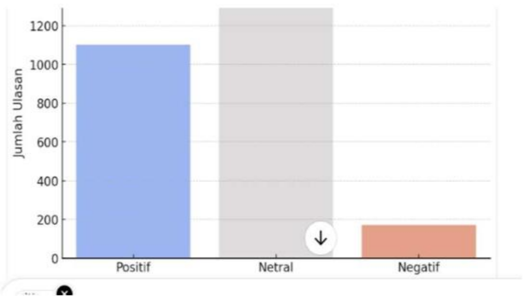

Figure 3. Distribution of Sentiment Category

This phenomenon underscores the importance of further analysis on neutral sentiment, as this category has the potential to provide additional insights into areas of service that can be improved to change customer perceptions to be more positive. Significant positive sentiment indicates that most customers are quite satisfied with the service provided. In contrast, low negative sentiment suggests that a small number of customers are facing problems or dissatisfaction that need to be addressed by the company immediately.

4.2 Sentiment Distribution Using TextBlob

Sentiment analysis using the TextBlob method produces a different pattern. The neutral category shows a much greater dominance than the lexicon method, with the number of reviews exceeding 2000. Positive reviews are drastically lower, below 500 reviews, while negative sentiment is barely visible in the results of this analysis in Figure 4 and Figure 5. This significant difference highlights the difference in sensitivity between the two sentiment analysis methods used in this study.

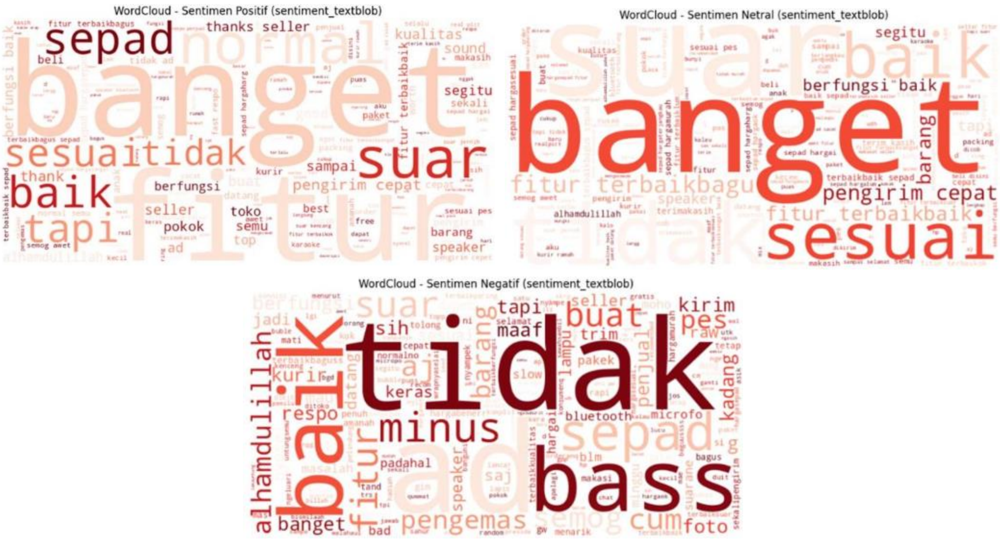

Figure 4. Word Cloud of Sentiment Analysis Using Texblob (a) Positive (b) Neutral (c) Negative

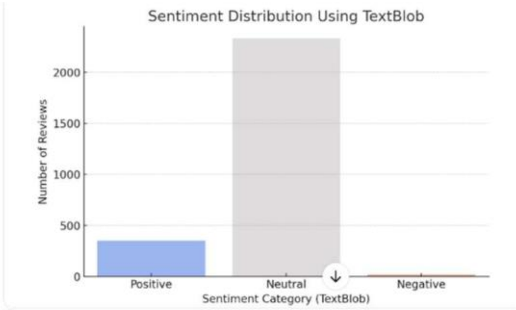

Figure 5. Distribution of Sentiment Category Using Texblob

The high number of neutral sentiments in the TextBlob method may reflect the characteristics of the algorithm which tends to be conservative in identifying sentiment polarity, resulting in more neutral categories than the lexicon method. Conversely, the low number of positive sentiments may indicate that TextBlob is more stringent in classifying positive reviews, so that many positive reviews detected by the lexicon method are not recognized by TextBlob.

4.3 Interpretation of Analysis Results

The difference in results between the Indonesian Lexicon and TextBlob methods reveals the importance of choosing the right sentiment analysis method. based on the business context and the type of data used. Lexicon methods that generate more positive categories may be more appropriate for use in business contexts that require high sensitivity to positive customer feedback. In contrast, the TextBlob method, which is more conservative in sentiment classification, may be more appropriate for businesses that want to accurately detect potential service improvements by focusing on reviews that truly show strong sentiment, both positive and negative.

The following Table 2 is the model evaluation matrix for 2 sentiment analysis methods.

Table 2. Proposed Model Performance

Category	Accuray	Precission	F1 Score
Texblob	63.54%,	53.72%,	56.34%,
Lexicon	78.52%,	78.62%,	68.11%,

The trend of neutral sentiment dominance in both analysis methods indicates that many customers have not fully experienced significant added value from the services offered. This finding encourages companies to explore more deeply certain aspects of their services that can be optimized to increase customer sentiment to be explicitly positive. This pattern also indicates that special attention must be paid to customer communication and interaction to positively impact customer perceptions.

Thus, the results of this study provide critical insights into how different sentiment analyses can be used in a complementary manner to gain a comprehensive picture of customer perceptions of a product or service, thereby enabling better and more strategic decision- making by companies in improving service quality.

5. CONCLUSION

This study has explored the effectiveness of sentiment analysis using Indonesian Lexicon and TextBlob based methods in business and e-commerce. The main results show that the distribution of user sentiment tends to be dominated by the neutral category in both analysis methods, with the TextBlob method showing a higher level of neutral sentiment than the Lexicon method. Meanwhile, the Lexicon method tends to be more sensitive in identifying positive reviews than TextBlob, which trends to be more conservative with an accuracy result of 78.52%.

The main finding of this study is that most customers give neutral reviews of the services or products offered. This indicates that there is room for service improvement to change customer perceptions in a more positive direction. In addition, the low proportion of negative sentiment suggests that customers are generally satisfied with the services provided. Still, there is potential that has not been fully utilized in improving customer experiences to be more explicitly positive.

The discussion in this study also emphasizes the importance of selecting a sentiment analysis method according to the company's specific needs. The Lexicon method proved more sensitive in capturing positive sentiment, while TextBlob was more stringent in sentiment classification. Therefore, companies should choose the method that best suits their analysis goals, whether to detect service improvement areas or confirm customer satisfaction.

Practically, this study provides valuable insights for businesses and e-commerce in utilizing sentiment analysis to develop more effective service strategies and increase customer satisfaction. In the future, it is recommended to conduct further research by combining lexicon-based approaches and machine learning in more depth to improve the accuracy and relevance of sentiment analysis results.

DATA AVAILABILITY

The dataset used in this study is publicly available on Kaggle.com.

REFERENCES

[1] S. Kumar, A. K. Kar, and P. V. Ilavarasan, "Applications of text mining in services management: A systematic literature review," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 1, p. 100008, 2021.

[2] M. Zhang, Y. Wang, and H. Olya, "Shaping social media analytics in the pursuit of organisational agility: a real options theory perspective," *Tour. Manag.*, vol. 88, p. 104415, 2022.

[3] E. Park, J. Kang, D. Choi, and J. Han, "Understanding customers' hotel revisiting behaviour: a sentiment analysis of online feedback reviews," *Curr. Issues Tour.*, vol. 23, no. 5, pp. 605–611, 2020.

[4] S. R. Nazar and T. Bhattasali, "Sentiment analysis of customer reviews," *Azerbaijan J. High Perform. Comput.*, vol. 4, no. 1, pp. 113–125, 2021.

[5] P. A. Bodhankar, R. K. Nasare, and G. Yenurkar, "Designing a sales prediction model in tourism industry and hotel recommendation based on hybrid recommendation," in *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*, IEEE, 2019, pp. 1224–1228.

[6] S. M. Ferdous, S. N. E. Newaz, S. B. S. Mugdha, and M. Uddin, "Sentiment Analysis in the Transformative Era of Machine Learning: A Comprehensive Review," *Stat. Optim. Inf. Comput.*, vol. 13, no. 1, pp. 331–346, 2025.

[7] S. Wong *et al.*, "EEG based automated seizure detection–A survey of medical professionals," *Epilepsy Behav.*, vol. 149, p. 109518, 2023.

[8] L.-Y. Ong and M.-C. Leow, "Comparative Study of Leveraging Big Data Processing Techniques for Sentiment Analysis," in *2023 10th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, 2023, pp. 343–348.

[9] I. Chalkias, K. Tzafilkou, D. Karapiperis, and C. Tjortjis, "Learning analytics on YouTube educational videos: exploring sentiment analysis methods and topic clustering," *Electronics*, vol. 12, no. 18, p. 3949, 2023.

[10] S. S. Sharma and G. Dutta, "SentiDraw: Using star ratings of reviews to develop domain specific sentiment lexicon for polarity determination," *Inf. Process. Manag.*, vol. 58, no. 1, p. 102412, 2021.

[11] S. Shayaa *et al.*, "Sentiment analysis of big data: methods, applications, and open challenges," *Ieee Access*, vol. 6, pp. 37807–37827, 2018.

[12] A. Taparia and T. Bagla, "Sentiment analysis: predicting product reviews' ratings using online customer reviews," *Available SSRN 3655308*, 2020.

[13] M. H. D. T. A. Bangsa, S. Priyanta, and Y. Suyanto, "Aspect-based sentiment analysis of online marketplace reviews using convolutional neural network," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 14, no. 2, pp. 123–134, 2020.

[14] R. Ahuja, A. Chug, S. Kohli, S. Gupta, and P. Ahuja, "The impact of features extraction on the sentiment analysis," *Procedia Comput. Sci.*, vol. 152, pp. 341–348, 2019.

[15] S. Verma, "Sentiment analysis of public services for smart society: Literature review and future research directions," *Gov. Inf. Q.*, vol. 39, no. 3, p. 101708, 2022.

[16] H. J. Lee, M. Lee, H. Lee, and R. A. Cruz, "Mining service quality feedback from social media: A computational analytics method," *Gov. Inf. Q.*, vol. 38, no. 2, p. 101571, 2021.

[17] A. Trivedi, S. Mondal, S. S. Mathew, and A. Jadhav, "Customer Review Analysis of Indian Multiplex Chain Services," in *2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, IEEE, 2024, pp. 1–6.

[18] M. Maggon and H. Chaudhry, "Exploring Relationships Between Customer Satisfaction and Customer Attitude from Customer Relationship Management Viewpoint: An Empirical Study of Leisure Travellers," *FIIB Bus. Rev.*, vol. 7, pp. 57–65, Mar. 2018, doi: 10.1177/2319714518766118.

[19] M. S. Hossain, M. F. Rahman, M. K. Uddin, and M. K. Hossain, "Customer sentiment analysis and prediction of halal restaurants using machine learning approaches," *J. Islam. Mark.*, vol. 14, no. 7, pp. 1859–1889, 2022.

[20] Z. Wang, V. Joo, C. Tong, X. Xin, and H. C. Chin, "Anomaly detection through enhanced sentiment analysis on social media data," in *2014 IEEE 6th international conference on cloud computing technology and science*, IEEE, 2014, pp. 917–922.

[21] A. Balahur, R. Mihalcea, and A. Montoyo, "Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications," *Computer Speech & Language*, vol. 28, no. 1. Elsevier, pp. 1–6, 2014.

[22] A. Mahgoub *et al.*, "Sentiment analysis: Amazon electronics reviews using bert and textblob," in *2022 20th International Conference on Language Engineering (ESOLEC)*, IEEE, 2022, pp. 6–10.

[23] K. Machová, M. Mikula, X. Gao, and M. Mach, "Lexicon-based sentiment analysis using the particle swarm optimization," *Electronics*, vol. 9, no. 8, p. 1317, 2020.

[24] C. S. G. Khoo and S. B. Johnkhan, "Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons," *J. Inf. Sci.*, vol. 44, no. 4, pp. 491–511, 2018.

[25] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexicon-based methods for sentiment analysis," *Comput. Linguist.*, vol. 37, no. 2, pp. 267–307, 2011.