Jurnal Krisnadana (Jurnal Komputer, Sistem Kendali & Jaringan)
Vol 5, Issue 3, May 2026, pp. 459-470
https://doi.org/10.58982/krisnadana.v5i3.1235

E-ISSN: 2808-3563
P-ISSN: 2808-3520

Integrating WebQual 4.0 and Machine Learning for Evaluating JDIH Tabanan Regency Website

Gusti Putu Sutrisna Wibawa[1*], Putu Sugiartawan[2], Ni Wayan Sumartini Saraswati[3]

[1*,2,3] Master of Informatics Department Institute of Business and Technology Indonesia, Denpasar, Indonesia
[1*]sutrisnawibawa@student.instiki.ac.id ; [2] putu.sugiartawan@instiki.ac.id ; [3] sumartini.saraswati@instiki.ac.id

ARTICLE INFO (8 pt)

Article history:
Received 16 May 2026
Revised 3 June 2026
Accepted 20 June 2026

Keywords:
WebQual 4.0;
sentiment analysis;
Random Forest;
E-Government;
Machine Learning;

ABSTRACT

The Legal Documentation and Information Network (JDIH) website of Tabanan Regency is a digital public service platform that provides access to regional regulations and legal documents. Evaluating website quality is important to ensure that the services provided are able to meet user needs and expectations. Previous studies on website evaluation generally applied either WebQual 4.0 or sentiment analysis separately, while studies integrating both approaches remain relatively limited. Therefore, this study combines WebQual 4.0 and machine learning–based sentiment analysis to provide a more comprehensive website evaluation. This study employed a mixed-methods approach involving 120 respondents and 120 user comments. Website quality evaluation was conducted using the WebQual 4.0 framework through three dimensions: usability, information quality, and service interaction. Sentiment analysis was performed using preprocessing techniques, TF-IDF feature extraction, and the Random Forest algorithm. The results showed that all WebQual dimensions achieved excellent categories, with usability obtaining the highest mean score of 4.487. Positive sentiment dominated user comments at 59.2%, while the Random Forest model achieved an accuracy rate of 79%. This integrated approach provided a more comprehensive evaluation and identified several technical aspects that still require improvement, including system performance, feature optimization, and interface consistency.

.

1. Introduction

As a digital public service platform, the Legal Documentation and Information Network (JDIH) website is expected to provide accurate, complete, relevant, and easily accessible legal information for users. In addition to information quality, usability aspects such as ease of use, navigation structure, interface appearance, and system responsiveness also play important roles in shaping user experience and satisfaction. However, several common issues are still frequently found in government websites, including slow system

performance, unclear navigation, outdated information, and inconsistent interface design. These issues may reduce user satisfaction and limit the effectiveness of digital public services.

Website quality evaluation is therefore essential to determine the extent to which a website meets user expectations and service standards. One of the most widely adopted approaches for website quality evaluation is WebQual 4.0, which measures website quality based on user perceptions through three main dimensions: usability, information quality, and service interaction. Previous studies have shown that WebQual is effective in evaluating website quality across various domains, including public service and e-government websites [1], [2], [4], [8]. In addition, several studies reported that WebQual provides a systematic framework for measuring website quality from the perspective of user experience and satisfaction [3], [5], [10].

Although WebQual offers a systematic framework for evaluating website quality, it mainly depends on quantitative questionnaire-based assessments and may not fully capture direct user opinions expressed through comments or reviews. Along with the rapid development of artificial intelligence (AI) and natural language processing (NLP), sentiment analysis has become an alternative approach for exploring user perceptions from textual data in a more comprehensive manner. Through sentiment analysis, user opinions can be automatically classified into positive, neutral, and negative sentiments [11], [22].

Recent studies have widely implemented machine learning techniques for sentiment classification because of their effectiveness in identifying patterns in textual data [16], [21], [24]. Among various machine learning algorithms, Random Forest has demonstrated strong classification performance, robustness, and the ability to reduce overfitting [19], [25]. Furthermore, TF-IDF (Term Frequency–Inverse Document Frequency) has been proven effective in transforming textual data into numerical representations for sentiment classification tasks [18].

Several previous studies have separately implemented WebQual and sentiment analysis for evaluating websites and digital services. Previous research evaluating the JDIH Tabanan Regency website using WebQual 4.0 and User Experience Questionnaire (UEQ) indicated that the website generally achieved positive usability and user experience results. However, the study mainly focused on structured questionnaire-based evaluation and did not analyze direct user opinions through machine learning–based sentiment analysis. More broadly, studies integrating quantitative website quality evaluation and machine learning–based sentiment analysis within a unified framework remain limited, particularly in the context of local government legal information websites.

Therefore, this study proposes an integrated approach combining WebQual 4.0 and machine learning–based sentiment analysis to provide a more comprehensive evaluation of website quality. The novelty of this study lies in integrating structured quantitative user assessments with direct textual user opinions within a single evaluation framework. Through this integration, the study not only evaluates website quality based on user perceptions but also identifies user feedback, complaints, and suggestions that may not be fully captured through questionnaires alone.

This study aims to evaluate the quality of the JDIH Tabanan Regency website using WebQual 4.0 and sentiment analysis based on the Random Forest algorithm. Data were

collected through questionnaires distributed to 120 respondents and 120 user comments obtained from website users. Sentiment analysis was conducted using preprocessing techniques, TF-IDF feature extraction, and the Random Forest classification algorithm. The integration of these approaches is expected to provide deeper insights into user perceptions and generate more comprehensive recommendations for improving e-government website quality.

2. Method

To obtain a more comprehensive evaluation of the JDIH Tabanan Regency website, this study combines WebQual 4.0 and machine learning–based sentiment analysis within a mixed-methods approach. WebQual 4.0 was utilized to quantitatively assess user perceptions of website quality through the dimensions of usability, information quality, and service interaction. In contrast, sentiment analysis was employed to explore direct user opinions expressed in textual comments, including feedback, suggestions, and complaints related to website performance and user experience.

The integration of both approaches allowed the quantitative results obtained from WebQual to be compared and supported by qualitative findings from sentiment analysis. This approach enabled the study to examine the consistency between user perception scores and user opinions while also identifying additional technical issues that may not be fully captured through questionnaire-based evaluation alone. As a result, the combined method provides a broader and more in-depth understanding of website quality by integrating quantitative user satisfaction measurements with qualitative user perspectives.

Using a purposive sampling technique, data were collected from users who met the research criteria through structured questionnaires and open-ended comments.

The research process began with data collection through questionnaires and user comments obtained from users of the JDIH Tabanan Regency website. The sampling technique used in this study was purposive sampling, where respondents were selected based on specific criteria relevant to the research objectives. Purposive sampling is commonly used in quantitative research to obtain respondents who meet specific research characteristics [26]. The respondents in this study were users who had previously accessed the website and were willing to provide assessments and feedback regarding their experience using the website. A total of 120 respondents participated in this study.

The overall research framework is illustrated in Figure 1. The process started with data collection and preparation, followed by two analytical approaches: WebQual 4.0 analysis and sentiment analysis. The results from both approaches were then integrated to produce a comprehensive evaluation of website quality.

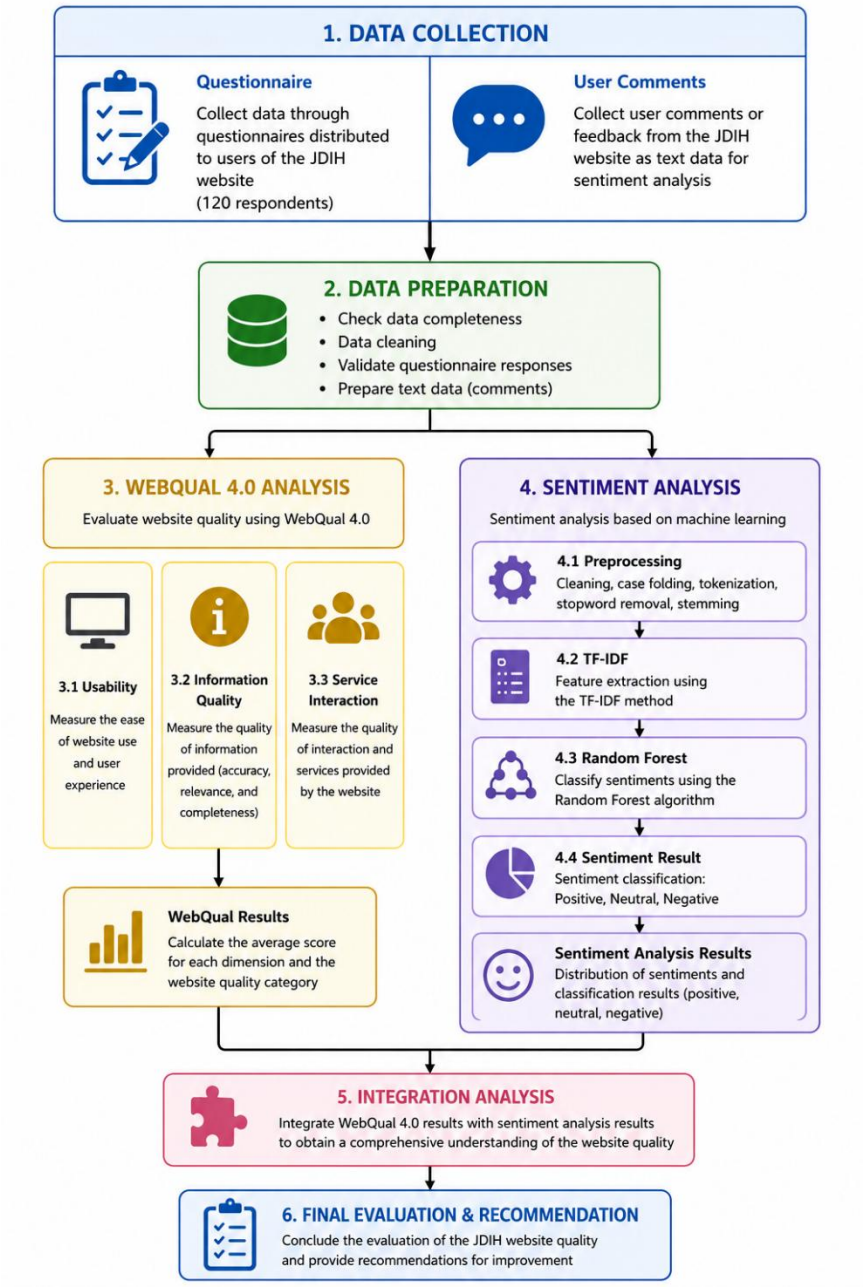

Figure 1. Research Flowchart

Website quality evaluation was conducted using the WebQual 4.0 method developed by Barnes and Vidgen [1], [2]. WebQual 4.0 is widely used to evaluate website quality based on user perceptions through three main dimensions: usability, information quality, and service interaction [3], [5]. The dimensions used in this study are presented in Table 1.

Table 1. WebQual 4.0 Dimensions

Dimension	Description
Usability	Measures ease of use, navigation, and interface appearance
Information Quality	Measures accuracy, relevance, clarity, and completeness of information
Service Interaction	Measures interaction quality, responsiveness, and user trust toward the website

Data collection for the WebQual analysis was conducted using questionnaires based on a five-point Likert scale ranging from strongly disagree to strongly agree. Likert scale is commonly used in quantitative research to measure respondent perceptions systematically [26]. The score of each indicator was calculated using the mean value formula:

$$\bar{X} = \frac{\sum X_i}{n} \quad (1)$$

where $\bar{X}$represents the mean score, $X_i$represents the respondent score, and $n$represents the total number of respondents.

In addition to questionnaires, this study also utilized user comments as textual data for sentiment analysis. Sentiment analysis is a text mining approach used to identify and classify opinions into sentiment categories such as positive, neutral, and negative [22].

Before classification, the textual data underwent several preprocessing stages to improve text quality and reduce noise. Text preprocessing is an important stage in natural language processing to improve classification performance [14]. The preprocessing stages used in this study are presented in Table 2.

Table 2. Text Preprocessing Stages

Step	Description
Cleaning	Removing punctuation, symbols, and unnecessary characters
Case folding	Converting all text into lowercase
Tokenization	Splitting sentences into individual words or tokens
Stopward Removal	Removing irrelevant and frequently occurring words
Stemming	Converting words into their root forms

After preprocessing, the text data were transformed into numerical representations using the TF-IDF (Term Frequency–Inverse Document Frequency) method. TF-IDF is widely used in text mining and information retrieval to measure the importance of words within documents [18].

The TF-IDF formula is defined as follows:

$$TF\text{-}IDF(t,d) = TF(t,d) \times IDF(t) \quad (2)$$

where $TF(t,d)$ represents the frequency of term $t$in document d, and $IDF(t)$represents the inverse document frequency.

The inverse document frequency is calculated using the following formula:

$$IDF(t) = \log \frac{N}{df(t)} \quad (3)$$

where N represents the total number of documents and $df(t)$represents the number of documents containing term t.

The sentiment classification process was performed using the Random Forest algorithm. Random Forest is an ensemble machine learning algorithm that combines multiple decision trees to improve classification accuracy and reduce overfitting [16], [19]. This algorithm is widely used in sentiment classification because of its good performance in handling textual data [11]. The model classified user comments into three sentiment categories: positive, neutral, and negative.

The dataset was divided into training and testing data using an 80:20 ratio. The training data were used to train the classification model, while the testing data were used to evaluate model performance

The implementation process was conducted using Python programming language with several supporting libraries, including NLTK for text preprocessing [14], Scikit-learn for TF-IDF transformation and machine learning classification [19], and Sastrawi for Indonesian stemming processes.

Model performance evaluation was conducted using confusion matrix and several evaluation metrics, namely accuracy, precision, recall, and F1-score [20], [21]. Accuracy measures the overall correctness of the model predictions.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision measures the proportion of correctly predicted positive instances.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Recall measures the model's ability to identify all positive instances.

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

F1-score represents the harmonic mean between precision and recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (7)$$

In this study, classification was performed using a multi-class approach consisting of positive, neutral, and negative sentiment classes. Confusion matrix was also used to visualize classification results and identify misclassification patterns among sentiment classes.

Finally, the results of the WebQual analysis and sentiment analysis were integrated to obtain a more comprehensive understanding of website quality. The integration process enabled the combination of quantitative user assessments and qualitative user opinions within a unified evaluation framework. The integrated findings were then used to identify the strengths and weaknesses of the JDIH Tabanan Regency website and to provide recommendations for future improvements.

3. Result and Discussions

The evaluation of the JDIH Tabanan Regency website was conducted using the WebQual 4.0 method based on three main dimensions, namely usability, information quality, and service interaction. The analysis results were obtained from questionnaire responses collected from 120 respondents. The results of the WebQual evaluation are presented in Table 3.

Table 3. WebQual 4.0 Evaluation Results

Dimension	Mean Score	Category
Usability	4.487	Excellent
Information quality	4.432	Excellent
Service interaction	4.398	Excellent

Based on Table 3, the usability dimension obtained the highest mean score of 4.487, indicating that users perceived the website as easy to use, easy to navigate, and visually understandable. Users also considered the interface structure and website features sufficiently user-friendly. This finding indicates that usability plays an important role in improving user experience and satisfaction when accessing e-government websites.

The information quality dimension obtained a mean score of 4.432, which also falls into the excellent category. This result indicates that users considered the information provided by the website to be accurate, relevant, understandable, and sufficiently complete. The availability of legal documents and regulatory information contributed positively to user perceptions regarding website quality.

Meanwhile, the service interaction dimension obtained a mean score of 4.398. Although categorized as excellent, several users still reported limitations related to system responsiveness and interaction quality. This finding suggests that improvements are still needed in several technical aspects to enhance user interaction and service quality.

Overall, the WebQual analysis results indicate that the JDIH Tabanan Regency website has a good level of quality based on user perceptions. These findings are consistent with previous studies showing that WebQual 4.0 is effective in evaluating website quality based on usability, information quality, and interaction quality [1], [2], [5].

In addition to quantitative evaluation using WebQual 4.0, this study also analyzed user opinions through sentiment analysis based on textual comments. The sentiment classification process categorized user comments into positive, neutral, and negative sentiments using the Random Forest algorithm. The sentiment distribution results are presented in Table 4.

Table 4. Sentiment Distribution Results

Sentiment	Total	Percentage
Positive	72	60.0%
Neutral	45	37.5%
Negative	3	2.5%

Based on Table 4, positive sentiment dominated the user comments with a percentage of 60.0%. This result indicates that most users had positive perceptions regarding the website quality, particularly in terms of ease of use, information accessibility, and usefulness of legal information services

Neutral sentiment accounted for 37.5% of the comments. Neutral comments generally contained suggestions, recommendations, or objective opinions regarding several website features without explicitly expressing positive or negative emotions.

Meanwhile, negative sentiment represented only 2.5% of the total comments. Negative opinions were primarily related to technical issues such as slow access speed, lack of feature optimization, and interface consistency. Although the percentage was relatively small, these comments provide important insights for future website improvement.

The performance of the Random Forest classification model was evaluated using accuracy, precision, recall, F1-score, and confusion matrix. The evaluation results are presented in Table 5.

Table 5. Random Forest Classification Results

Category	Precision	Recall	F1-Score	Support
Neutral	0.67	0.89	0.76	9
Positive	0.92	0.73	0.81	15
Accuracy	-	-	0.79	24

Based on Table 5, the Random Forest model achieved an accuracy of 0.79 or 79%, indicating that the model was able to classify sentiment categories reasonably well. The positive sentiment category obtained the highest precision value of 0.92, indicating that the model had a high level of correctness in predicting positive comments.

Meanwhile, the neutral category obtained a recall value of 0.89, indicating that the model was capable of identifying most neutral comments correctly. However, several classification errors still occurred between positive and neutral sentiments due to similarities in sentence patterns and vocabulary usage.

The confusion matrix results are presented in Table 6

Table 6. Confusion Matrix Results

Actual/Predicted	Neutral	Positive
Neutral	8	1
Positive	4	11

Based on Table 6, the model correctly classified 8 neutral comments and 11 positive comments. However, one neutral comment was incorrectly classified as positive, while four positive comments were classified as neutral. These results indicate that the Random Forest model was able to recognize sentiment patterns relatively well, although several overlapping characteristics between sentiment categories still caused classification errors.

The negative sentiment category did not appear in the testing data due to the limited number of negative comments within the dataset. This condition indicates that the dataset still experienced class imbalance, where positive comments dominated the overall data distribution.

Overall, the evaluation results demonstrate that Random Forest provided satisfactory performance for sentiment classification in this study. These findings are consistent with previous studies indicating that Random Forest performs effectively in text classification and sentiment analysis tasks [11], [19], [25].

The integration of WebQual 4.0 and sentiment analysis provided a more comprehensive evaluation of the JDIH Tabanan Regency website. The WebQual results demonstrated that all dimensions achieved excellent category scores, while the sentiment analysis results showed that positive sentiment dominated user comments.

The relationship between the WebQual 4.0 evaluation results and sentiment analysis findings demonstrates a consistent pattern regarding user perceptions of the JDIH Tabanan Regency website. The high usability score of 4.487 was aligned with the dominance of positive sentiment in user comments. Many users expressed positive opinions regarding ease of navigation, interface simplicity, and accessibility of legal information services. This consistency indicates that usability is one of the primary factors contributing to positive user experience and satisfaction.

Similarly, the information quality dimension achieved a high mean score of 4.432, indicating that users perceived the information available on the website as accurate, relevant, and understandable. These findings were reinforced by sentiment analysis results, where positive comments frequently highlighted the completeness and usefulness of legal documents and regulatory information provided by the website.

Although the service interaction dimension also obtained an excellent category score of 4.398, sentiment analysis revealed several technical concerns that were not fully reflected in the questionnaire-based evaluation. Some users reported issues related to loading speed, feature optimization, interface consistency, and system responsiveness. This finding

suggests that quantitative evaluation through WebQual may not fully capture detailed technical experiences encountered by users during website interaction.

In addition, sentiment analysis identified several aspects that were not explicitly represented within the WebQual dimensions, such as mobile accessibility issues, broken links, and website performance stability. Therefore, sentiment analysis complements the WebQual evaluation by providing deeper qualitative insights into user experiences, complaints, and suggestions.

These findings indicate that integrating WebQual 4.0 and machine learning–based sentiment analysis not only strengthens the validity of website evaluation results but also enables a more comprehensive understanding of website strengths and weaknesses. The integrated approach allows quantitative user perception measurements to be supported by direct textual user opinions, thereby producing more detailed recommendations for improving e-government website quality and digital public services.

However, sentiment analysis also revealed several issues that were not fully captured through questionnaire-based evaluation. Some users reported technical issues related to system performance, feature completeness, loading speed, and interface consistency. Although these issues did not significantly reduce the overall WebQual scores, they provide important insights regarding aspects requiring further improvement.

These findings demonstrate that integrating WebQual and machine learning–based sentiment analysis can provide a deeper understanding of website quality by combining structured user assessments and direct user opinions. Therefore, this integrated approach can serve as an effective evaluation framework for e-government websites and digital public service platforms.

The findings of this study are consistent with previous research emphasizing the importance of usability and information quality in improving user satisfaction [1], [2], [10]. However, this study extends previous work by integrating sentiment analysis to complement quantitative evaluation results with direct user opinions extracted from textual comments.

4. Conclusion

This study evaluated the quality of the JDIH Tabanan Regency website using an integrated approach combining WebQual 4.0 and machine learning–based sentiment analysis. The WebQual evaluation results showed that all dimensions, namely usability, information quality, and service interaction, achieved excellent category scores based on user perceptions. Among these dimensions, usability obtained the highest mean score, indicating that users perceived the website as easy to use and navigate. The sentiment analysis results demonstrated that positive sentiment dominated user comments, indicating that most users had favorable perceptions regarding the website quality and legal information services provided. The Random Forest classification model achieved an accuracy of 79%, showing satisfactory performance in classifying user sentiments into positive and neutral categories. However, the negative sentiment category was limited due to the imbalance of sentiment distribution within the dataset. The integration of WebQual

4.0 and sentiment analysis provided a more comprehensive evaluation framework by combining quantitative user assessments and qualitative user opinions. The integrated findings revealed that, although the website generally received positive evaluations, several technical aspects such as system responsiveness, feature optimization, and interface consistency still require improvement. Overall, this study demonstrates that integrating WebQual and machine learning–based sentiment analysis can provide deeper insights into user perceptions and website quality evaluation. The proposed approach can also be applied as an alternative evaluation framework for e-government and digital public service websites. Future studies are recommended to use larger and more balanced datasets as well as more advanced machine learning or deep learning methods to improve sentiment classification performance.

Acknowledgment

The authors would like to acknowledge the important contributions of the respondents whose input informed this research. We also extend our special thanks to the Legal Department of the Tabanan Regency Secretariat, as the administrator of the Tabanan Regency JDIH, for supporting our research efforts and providing essential resources. Finally, we extend our appreciation to all those who assisted in this project and contributed to its completion in various capacities..

References

[1] A. D. Putri, R. Mahendra, and S. Wijaya, "Analisis kualitas website menggunakan WebQual 4.0 pada layanan publik," *Jurnal Sistem Informasi*, vol. 12, no. 1, pp. 45–55, 2025.

[2] K. Huda, M. F. Rahman, and A. Prasetyo, "Website quality evaluation using WebQual 4.0 method," *International Journal of Information Systems*, vol. 10, no. 2, pp. 120–130, 2025.

[3] L. Rachmayani, D. Purnama, and Y. Saputra, "Website quality analysis using WebQual approach," *Jurnal Teknologi Informasi*, vol. 9, no. 1, pp. 33–42, 2025.

[4] E. Meidika, N. Kurniawan, and R. Aditya, "Evaluation of government website quality using WebQual," *Journal of E-Government Studies*, vol. 8, no. 2, pp. 88–98, 2024.

[5] D. H. Pamungkas and Subhan, "Analysis of website quality using WebQual and IPA method," *Jurnal Informatika*, vol. 11, no. 1, pp. 67–75, 2023.

[6] E. D. Yang, H. Lee, and J. Kim, "WebQual evaluation in e-learning systems," *International Journal of Education Technology*, vol. 7, no. 2, pp. 50–60, 2024.

[7] W. Murdoko, "Website evaluation using WebQual method," *Jurnal Sistem Komputer*, vol. 6, no. 1, pp. 22–30, 2023.

[8] A. K. Wijayanti, F. Nugroho, and I. Saputri, "E-government website evaluation using WebQual," *Journal of Digital Government*, vol. 5, no. 2, pp. 77–85, 2024.

[9] W. Kusumawardhana, "Literature review of WebQual method," *Jurnal Teknologi dan Informasi*, vol. 13, no. 1, pp. 10–20, 2025.

[10] M. A. Nurrahman, "User satisfaction analysis using WebQual 4.0," *Jurnal Sistem Informasi*, vol. 14, no. 1, pp. 55–65, 2025.

[11] M. T. Tusar and M. T. Islam, "Sentiment analysis using machine learning techniques," *International Journal of Computer Applications*, vol. 174, no. 5, pp. 15–22, 2021.
[12] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the NAACL-HLT Conference*, Minneapolis, MN, USA, 2019, pp. 4171–4186.
[13] Y. Goldberg, *Neural Network Methods in Natural Language Processing*. Cambridge, U.K.: Cambridge University Press, 2022.
[14] S. Bird, E. Klein, and E. Loper, *Natural Language Processing with Python*. Sebastopol, CA, USA: O'Reilly Media, 2021.
[15] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2021.
[16] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 4th ed. Burlington, MA, USA: Morgan Kaufmann, 2021.
[17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2022.
[18] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge University Press, 2021.
[19] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022.
[20] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2021.
[21] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2022.
[22] B. Liu, *Sentiment Analysis and Opinion Mining*. San Rafael, CA, USA: Morgan & Claypool Publishers, 2021.
[23] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken, NJ, USA: Pearson, 2021.
[24] F. Chollet, *Deep Learning with Python*, 2nd ed. Shelter Island, NY, USA: Manning Publications, 2022.
[25] J. Brownlee, *Machine Learning Mastery with Python*. Melbourne, Australia: Machine Learning Mastery, 2023.
[26] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung, Indonesia: Alfabeta, 2022.