Multidisciplinary Output Research For Actual and International Issue (MORFAI Journal) ISSN (e): 2808-6635
Volumes 6 No. 4 (2026)

SENTIMENT ANALYSIS OF THE REPUTATION OF THE GLOBAL TRAINING AND TEST CENTER (GTTC) SUKABUMI

Taufik Nurhadi[1]*, Hesri Mintawati[2], Yusuf Iskandar[3]
Universitas Nusa Putra / Sukabumi
Universitas Nusa Putra / Sukabumi
Universitas Nusa Putra / Sukabumi
E-mail: Taufik.nurhadi@nusaputra.ac.id[1]*, hesri.mintawati@nusaputra.ac.id[2], yusuf.iskandar@nusaputra.ac.id[3]

Received	: 01 March 2026	Accepted	: 01 April 2026
Revised	: 15 March 2026	Published	: 21 May 2026

Abstract

The reputation of a global training institution is an important factor in increasing competitiveness and public trust in the quality of services provided. The Global Training and Test Center (GTTC) Sukabumi, as a global-based training and certification institution, has various sources of feedback from participants, partners, and stakeholders, most of which are in the form of unstructured text data. This study aims to analyze sentiment towards the reputation of the Global Training and Test Center (GTTC) Sukabumi using a text mining-based sentiment analysis approach. Research data were obtained from participant reviews, testimonials, and stakeholder comments collected through interviews and documentation. The method used is sentiment analysis with the help of Orange Data Mining software to classify sentiment into positive, negative, and neutral, and identify key topics that influence the institution's reputation. The results of this study are expected to provide an overview of public perception towards GTTC and become a basis for strategic decision-making to improve the quality of service and the institution's global reputation.

Keywords: sentiment analysis; institutional reputation; text mining; GTTC; Orange Data Mining

INTRODUCTION

An institution's reputation is a strategic asset that plays a crucial role in determining the sustainability and competitiveness of an organization, particularly globally oriented training and certification institutions. In today's era of globalization and digitalization, reputation is shaped not only by the quality of services provided but also by public perception, which is widely disseminated through various media, both in person and digitally. This perception serves as a crucial indicator in assessing the credibility, professionalism, and level of public trust in an institution. The Global Training and Test Center (GTTC) Sukabumi is an institution specializing in global standards-based training and certification. In carrying out its role, GTTC focuses not only on improving participant competency but also strives to build its image and reputation as a professional, trustworthy, and internationally competitive institution. However, in practice, an institution's reputation cannot always be measured directly, as it is largely shaped by the opinions, experiences, and subjective perceptions of stakeholders, such as training participants, partners, and the general public.

With the increasing use of information technology, various forms of feedback on GTTC services are widely available in the form of unstructured text data, such as comments, reviews, testimonials, and interview results. This data contains extremely rich and valuable information, but it is often underutilized due to the limitations of conventional analysis methods, which tend to be manual and less able to capture the hidden meanings within the text. In this context, a text mining-based sentiment analysis approach is a relevant and effective solution for processing text data into more structured and meaningful information. Sentiment analysis allows researchers to identify trends in public opinion toward an object, whether positive, negative, or neutral, while simultaneously understanding the factors influencing the formation of these perceptions. Thus, this approach not only provides a quantitative overview of sentiment distribution but also reveals deeper insights into stakeholder experiences and expectations. The use of data analysis technologies like Orange Data Mining further simplifies the process of exploring text data, offering a variety of features that support preprocessing, sentiment classification, and intuitive data visualization. This enables more systematic, efficient, and data-driven research without requiring complex programming skills.

Several previous studies have shown that sentiment analysis has been widely used in various fields, such as educational service evaluation, customer satisfaction analysis, and organizational reputation measurement. These studies demonstrate that this approach can provide a more objective and comprehensive picture than conventional methods. However, research specifically examining the reputation of global training institutions, particularly in a local context like GTTC Sukabumi, is still relatively limited. Based on these conditions, there is a research gap in utilizing text mining technology to analyze the reputation of training institutions in a more in-depth, data-driven manner. Therefore, this study aims to fill this gap by examining public sentiment toward the reputation of the Global Training and Test Center (GTTC) Sukabumi using a sentiment analysis approach. This research is expected to provide both theoretical and practical contributions. Theoretically, this research enriches the study of the use of text mining and sentiment analysis in the context of institutional reputation management. Practically, the results can serve as a basis for GTTC management in formulating strategies to improve service quality, strengthen the institution's image, and sustainably increase stakeholder satisfaction and trust.

Based on this background, the problem formulation in this research is as follows:

1. What are the results of the sentiment analysis on the reputation of the Sukabumi Global Training and Test Center (GTTC)?
2. How much do positive, negative, and neutral sentiments dominate in shaping GTTC's reputation?
3. What aspects or topics have the most influence on GTTC's reputation based on text data analysis?

LITERATURE REVIEW

In this study, the literature review focused on the concept of institutional reputation, the use of text mining, and the use of sentiment analysis to analyze public perception of an organization. This study provides an important foundation for understanding how text data can be processed into strategically valuable information to support decision-making. An organization's reputation is a collective perception formed from the experiences, interactions, and information received by the public about an institution. In recent years, research has shown that reputation is influenced not only by service quality but also by public opinion that develops through digital media and online communication platforms. Putri and Santoso (2022) state that organizational reputation is increasingly influenced by widely disseminated user reviews and testimonials, which can significantly shape an institution's image.

The development of information technology has led to an increasing volume of unstructured data, particularly in text form such as comments, reviews, and testimonials. In this context, text mining is an effective approach for extracting information from this data. According to Pratama et al. (2023), text mining can identify patterns, relationships between words, and key themes in text data sets, thus helping organizations understand public perception more deeply. A crucial stage in text mining is text preprocessing, which cleans data from noise and prepares it for analysis. This process includes tokenization, text normalization, stopword removal, and stemming. Rachmawati and Nugroho (2021) explain that the quality of preprocessing significantly determines the final analysis results, as errors at this stage can impact the overall accuracy of the model.

Furthermore, sentiment analysis is a technique used to identify and classify individual opinions or feelings toward an object. In recent years, the use of sentiment analysis has grown and is widely applied in various fields, such as business, education, and public services. According to Hidayat et al. (2024), sentiment analysis can provide a more objective picture of user satisfaction levels and perceptions compared to conventional methods. In practice, lexicon-based methods such as the Valence Aware Dictionary and Sentiment Reasoner (VADER) remain widely used due to their ease and efficiency in sentiment classification. Sari and Wijaya (2022) stated that VADER is effective in detecting sentiment polarity, especially in informal text data derived from user interactions.Furthermore, the topic modeling approach is also an important part of text data analysis. Topic modeling is used to identify key themes that emerge within a collection of documents. According to Kurniawan et al. (2023), this method can help researchers understand the focus of a discussion without having to manually read all the data.

One method commonly used in topic modeling is Latent Dirichlet Allocation (LDA). This method allows for grouping documents into several topics based on word distribution. Rahman and Putra (2021) explain that LDA is effective in identifying hidden structures in text data and is frequently used in document analysis-based research. Recent research also shows that combining sentiment analysis and topic modeling can produce a more comprehensive analysis. This approach not only identifies sentiment trends but also explains the factors that influence the formation of these opinions. This finding is supported by research by Anwar et al. (2024), which states that integrating these two methods can provide organizations with deeper insights into understanding stakeholder needs and perceptions. However, research specifically examining the reputation of global training institutions is still

relatively limited, particularly in the local context. Most previous studies have focused on the general business sector or formal education. Therefore, this study aims to fill this gap by analyzing the reputation of the Sukabumi Global Training and Test Center (GTTC) using a text mining-based sentiment analysis approach. Thus, this research is expected to contribute to the development of data-based reputation analysis studies, as well as serve as a reference for further research that examines public perceptions of training and certification institutions.

METHOD

The method used in this study is *text mining*- based *sentiment analysis* with the help of *Orange* Data Mining software on qualitative data in the form of comments, testimonials, and reviews related to the reputation of *the Global Training and Test Center* (GTTC) Sukabumi.

Text Mining

Text mining is the process of extracting information from unstructured text data to discover patterns, relationships, and hidden meanings. In this study, text mining was used to process data in the form of comments and testimonials into data that can be analyzed systematically. Pratama et al. (2023) stated that text mining can identify linguistic patterns and key themes in a collection of documents, making it highly relevant for analyzing public perception of an institution .

Text Preprocessing

Text preprocessing is the initial stage in text data analysis which aims to clean the data from noise so that it is more structured (Rani & Arora, 2016).
These stages include:

- Case folding (changing letters to lower case)
- Tokenization (word splitting)
- Stopword removal (removing common words)
- Stemming/Lemmatization

Sentiment Analysis

Sentiment analysis is a technique used to identify and classify opinions or emotions in text into positive, negative, and neutral categories. In this study, this method was used to understand stakeholders' perceptions of the reputation of GTTC Sukabumi. According to Sari et al. (2023), sentiment analysis can provide a more objective picture of user satisfaction levels than conventional methods because it can automatically process large amounts of data.

Valence Aware Dictionary and Sentiment Reasoner (VADER)

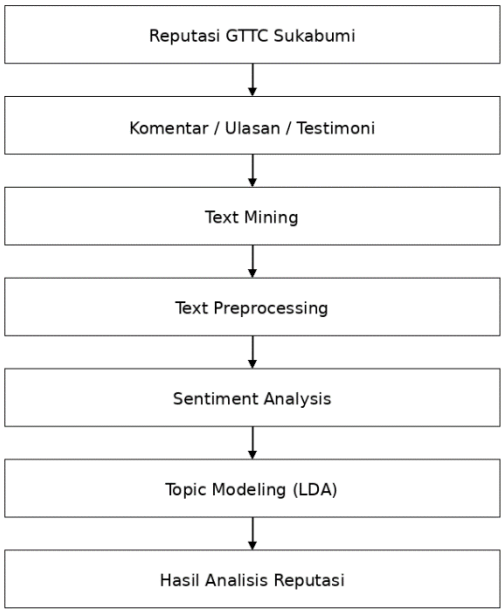
Valence Aware Dictionary and Sentiment Reasoner (VADER) is a lexicon-based sentiment analysis method used to determine sentiment polarity in text. This method works by assigning scores to words based on their emotional level, then accumulating these scores to determine sentiment categories. Wijaya and Putra (2022) stated that VADER is effective in classifying sentiment, especially in informal text data derived from user interactions.

Topic Modeling

Topic modeling is a technique used to identify key themes in a text dataset. This method works by grouping frequently occurring words into a specific topic. Kurniawan et al. (2023) explain that topic modeling is highly effective in analyzing large amounts of text data because it can automatically uncover hidden information patterns.

LDA (Latent Dirichlet Allocation)

Latent Dirichlet Allocation (LDA) is a topic modeling method used to group documents into topics based on word distribution. LDA assumes that each document consists of a combination of several latent topics. Rahman and Putra (2021) stated that LDA is effective in identifying hidden structures in text data and is frequently used in document analysis-based research.



Stages of Data Analysis Method:

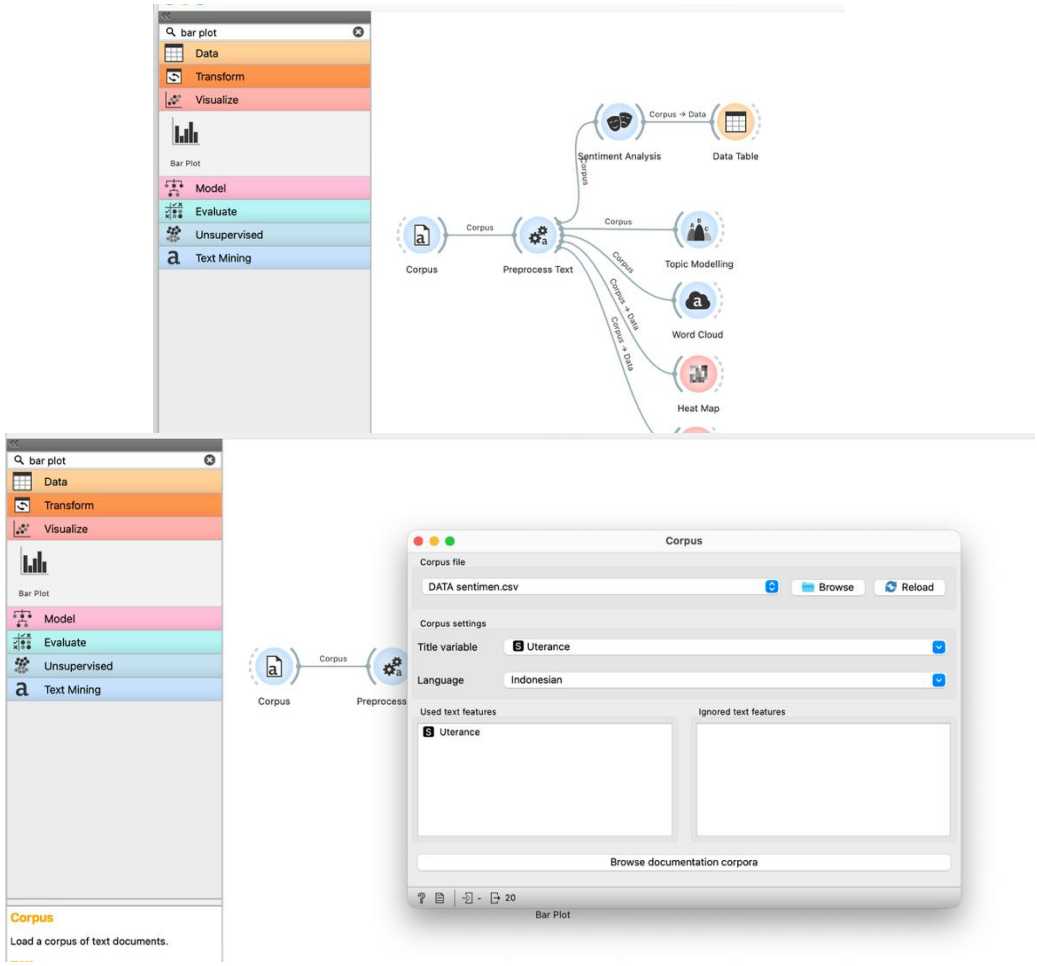
a. Institutional Reputation (GTTC)
   Identifying stakeholder perceptions of the quality of services, facilities, and professionalism of GTTC as a global training institution.
b. Stakeholder Comments
   Collecting data in the form of reviews, testimonials, and comments from training participants, partners, and other stakeholders.
c. Sentiment Analysis
   Classify text data into positive, negative, and neutral sentiment categories to determine public opinion trends.

Stages of Data Processing Method:

a. Interview
   Data was collected through interviews and documentation in the form of comments and testimonials relevant to GTTC's reputation. According to Anwar et al. (2024), collecting data based on direct experiences and user opinions is crucial for generating a more contextual and in-depth analysis.
b. Orange Data Mining
   Orange Data Mining is used as an analysis tool due to its ease of use and data visualization capabilities. This software enables text preprocessing, sentiment classification, topic modeling, and data visualizations such as word clouds and heatmaps. Prasetyo et al. (2022) state that using visual-based tools like Orange can improve analysis efficiency and simplify the interpretation of research results.

RESULTS AND DISCUSSION

Figure 2. ***Setiment Analysis*** **and** ***Topic Modeling***



Source: Orange Data Mining Software

Research Scenario

a. *Corpus* (Text Data Source)
In the initial stage, comment or review data is entered into the system using the Corpus widget in the Orange Data Mining application. Based on the display shown, the data file used is DATA sentiment.csv, which contains a collection of text in the form of stakeholder comments on the services of the Global Training and Test Center (GTTC) Sukabumi. In the Corpus settings, the text variable is defined in the Utterance section, which serves as the main source of data to be analyzed. In addition, the language used is set to Indonesian so that the system can process text more accurately and relevant to the research context.

a. *Text Preprocessing*
After the data has been successfully entered, the next step is to perform text preprocessing using the Preprocess Text widget. This step aims to clean the data of irrelevant elements and prepare it for better structure and analysis.
Based on the display shown, some of the preprocessing processes used include:

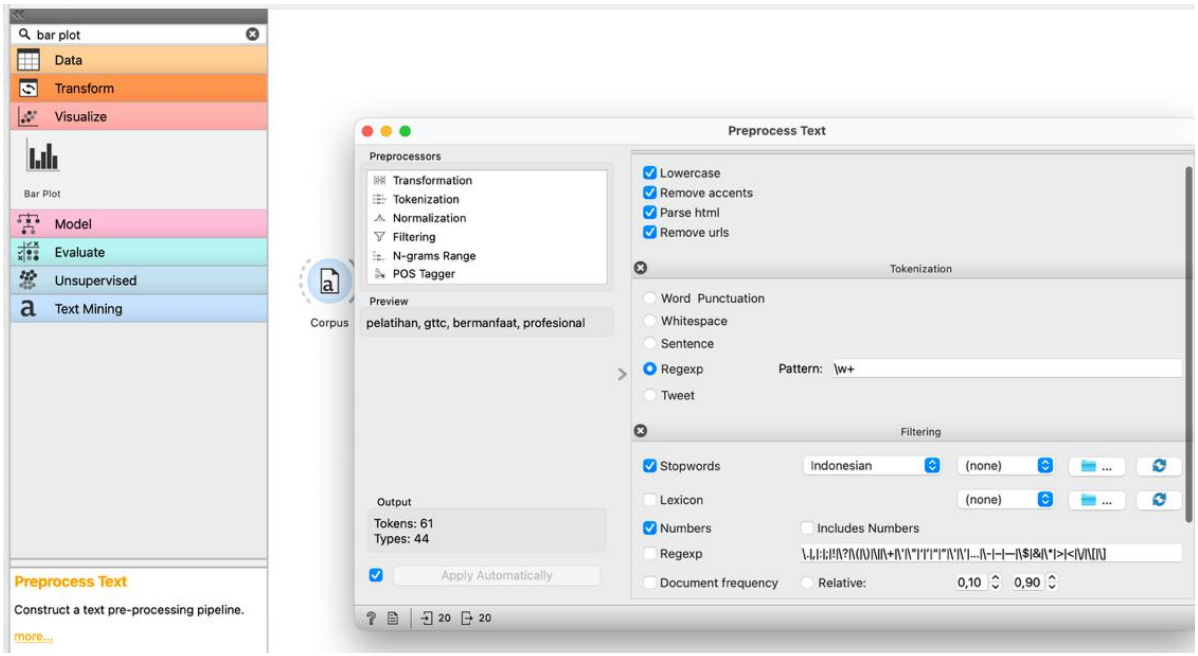
- Lowercase: Changes all text to lowercase to avoid spelling differences.
- Remove accents: Removes special characters or unnecessary accents.
- Parse HTML: Removes HTML elements if they are present in the data
- Remove URLs: Removes irrelevant links in the analysis

In the tokenization stage, the RegExp method is used, which allows word separation based on specific patterns, resulting in more accurate tokens. Additionally, filtering processes are performed, such as:

- Stopwords removal (Bahasa Indonesia) to remove common words that have no significant meaning
- Removal of numbers to avoid interference in text analysis

The results of this preprocessing process can be seen in the Preview section, which displays key words like "training," "gttc," "beneficial," and "professional." This indicates that the data has been successfully cleaned, leaving only words with high information value.

Figure 4. Preprocess Text

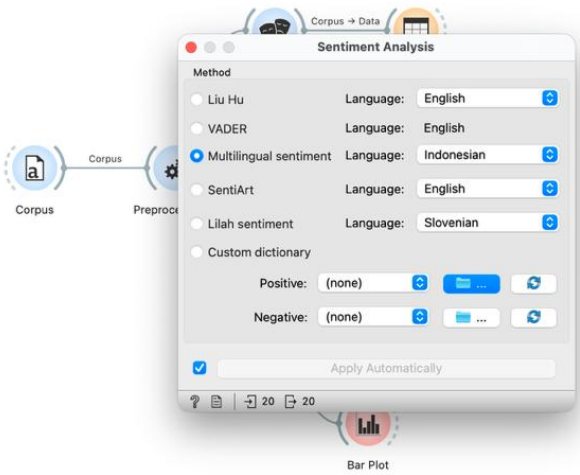


Source: Orange Data Mining Software

b. *Sentiment Analysis*

The sentiment analysis widget is used to identify and classify emotional expressions or attitudes contained in text into positive, negative, and neutral categories. In the Orange Data Mining application, this process generally utilizes a built-in English-based model. Therefore, if the data being analyzed is in Indonesian, additional adjustments are required, such as manual labeling or the use of a model that supports that language for more accurate analysis results. After going through the preprocessing stage, the text data is then analyzed using this widget, and the sentiment classification results are displayed systematically and structured through the data table feature.

Figure 5. Sentiment Analysis



Source: Orange Data Mining Software

c. Data Table

In the sentiment analysis stage, this study utilized the Sentiment Analysis feature available in the Orange Data Mining application. Based on the configuration used, the Multilingual Sentiment method was applied with Indonesian language settings. This method was chosen based on the characteristics of the data, which is in the form of Indonesian text, requiring a model capable of understanding the linguistic context more precisely. With the Multilingual Sentiment approach, the system can automatically classify text into sentiment categories without requiring the manual use of a dictionary or lexicon.

Figure 6. Data Table

	include title	Uterance True True	ID	sentiment
1		Pelatihan di GTTC sangat bermanfaat dan profesional	1	25
2		Fasilitas cukup baik namun perlu peningkatan	2	100
3		Saya sangat puas dengan pelayanan yang diberikan	3	50
4		Pelatihannya bagus tapi jadwal kurang fleksibel	4	50
5		Tidak puas dengan pelayanan administrasi	5	33.3333
6		Materi pelatihan sangat relevan dengan kebutuhan kerja	6	20
7		Kurang maksimal dalam penyampaian materi	7	0
8		Instruktur sangat kompeten dan ramah	8	33.3333
9		Fasilitas kurang lengkap	9	0
10		Sangat direkomendasikan untuk pelatihan global	10	33.3333
11		Pengalaman belajar yang menyenangkan	11	33.3333
12		Pelayanan lambat saat registrasi	12	-33.3333
13		Program pelatihan sangat membantu karir saya	13	25
14		Kurang puas dengan sistem pembelajaran	14	33.3333
15		GTTC memiliki kualitas yang baik	15	0
16		Materi terlalu cepat disampaikan	16	50
17		Lingkungan belajar nyaman dan kondusif	17	25
18		Kurang interaktif dalam pembelajaran	18	0
19		Sangat puas dengan hasil pelatihan	19	33.3333
20		Perlu peningkatan fasilitas teknologi	20	66.6667

Source: Orange Data Mining Software

Based on the results of processing using the Sentiment Analysis feature in the Orange Data Mining application, the output is a table displaying each comment along with its resulting sentiment value. The table consists of several main columns:

Utterance: contains the text of the respondent's comments

ID: shows the data identity number

Sentiment: the numerical value of the sentiment analysis results

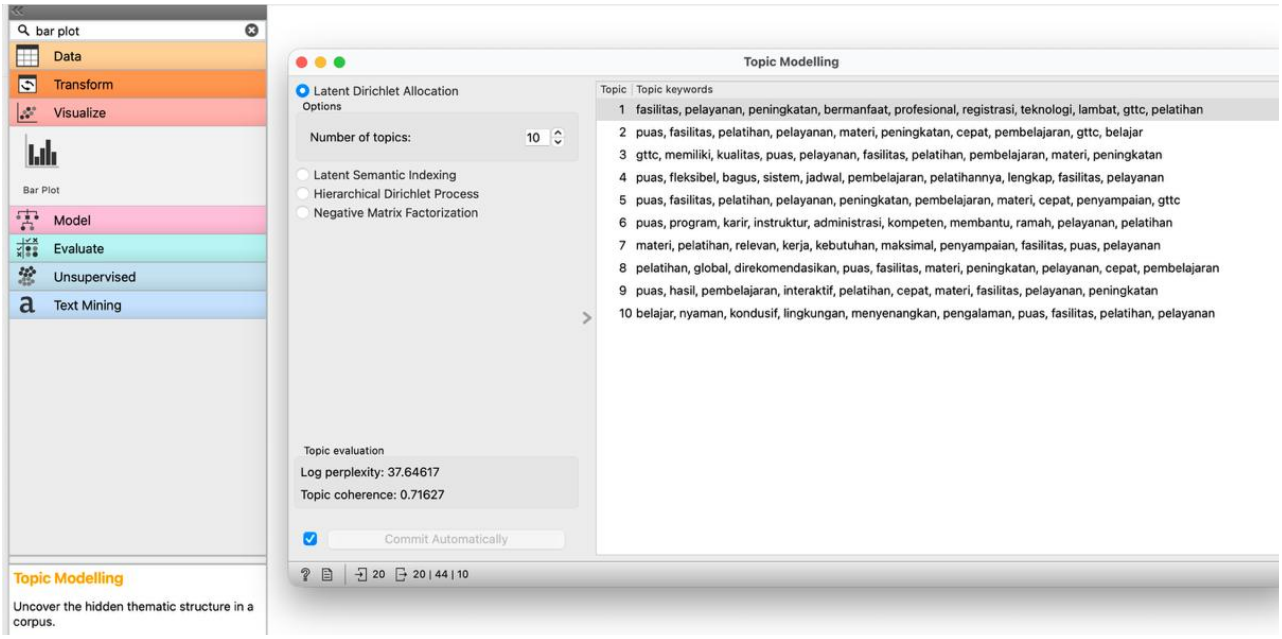
The sentiment column values are displayed as a numerical score representing the emotional level of each comment. This score can then be interpreted as a positive, neutral, or negative sentiment category.

d. Topic Modeling

Topic modeling is a feature used to identify key themes hidden within a collection of text. Using algorithms such as Latent Dirichlet Allocation (LDA), this feature groups frequently occurring words into specific

topics. The result is a list of topics along with keywords representing the content of each topic, making it easier to understand conversation patterns and the main focus of the information, especially when dealing with large and diverse data sets.

Figure 7. Topic Modeling



Source: Data Mining People Software

At this stage, the analysis was carried out using the Latent Dirichlet Allocation (LDA) method through the Orange Data Mining application to identify the main topics hidden in the text data set.

Based on the displayed data, the number of topics used in this study was 10. This number was chosen to obtain a more diverse representation of the themes from the overall comment data analyzed.

In addition, the model evaluation results show the values:

- Log Perplexity is 37.64617
- Topic Coherence of 0.71627

A relatively high topic coherence value indicates that the words in each topic are quite well related, so that the modeling results can be considered quite representative in describing the theme structure in the data.

e. *Word Cloud*

Word clouds are a simple yet informative visualization method for text analysis. This technique displays the most frequently occurring words from data such as interview results, comments, or feedback in the form of a word cloud image. This display allows users to quickly understand the main themes or topics most frequently discussed by respondents without having to read each sentence in detail.

Figure 8. Word Cloud

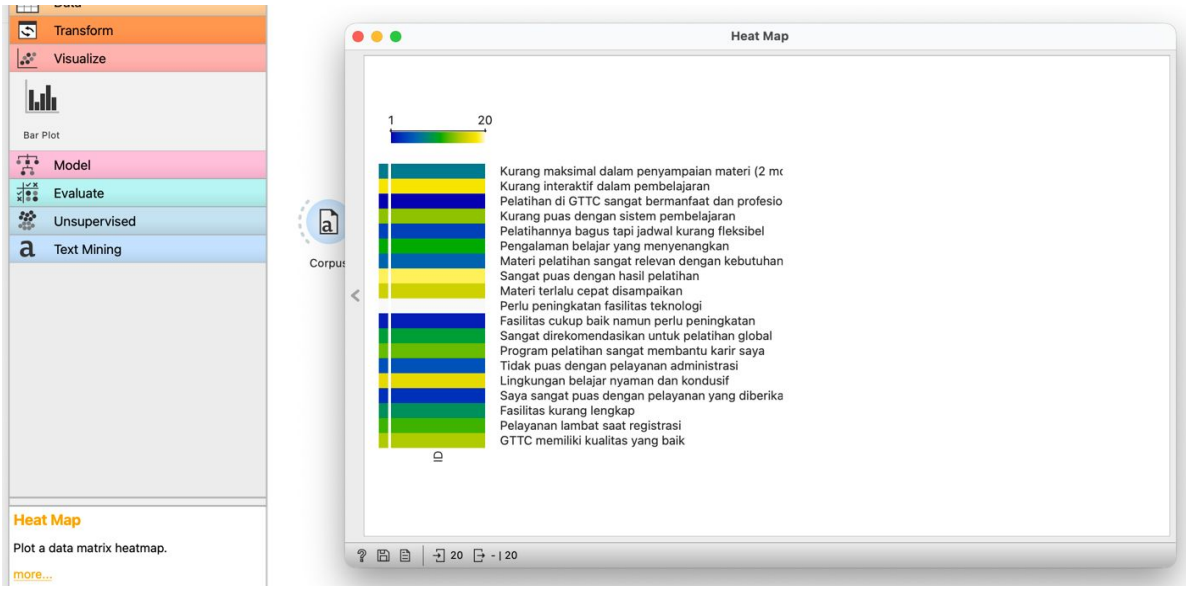
Source: Orange Data Mining Software

The word cloud visualization displays the frequency of word occurrences from text data (most likely survey results or training participant feedback). The larger the word size, the more frequently it appears in the dataset. The image shows that words such as "training," "satisfied," "facilities," "service," and "material" are the most dominant. This indicates that the main topic in the data relates to participants' experiences with the training, specifically aspects of satisfaction, the quality of the facilities, and the services provided. Supporting words such as "instructor," "schedule," "flexible," and "professional" also appear, indicating important factors influencing participants' perceptions. Overall, this word cloud provides a quick overview of the main focus of respondents' opinions.

f. *Heat Map*

A heat map is a form of data visualization used to illustrate the level of perception intensity based on text analysis using Orange Data Mining. The color variations displayed, ranging from blue representing low intensity to yellow indicating high intensity, make it easier to identify the most dominant, most frequently occurring, or significant statements in the classification and topic analysis process. This visualization facilitates a faster, more systematic, and more in-depth understanding of data patterns.

Figure 9. Heat Map



Source: Orange Data Mining Software

Heat maps are used to show the intensity or trend of values through color. In this context, each row represents a statement or review, while color indicates a specific level (e.g., sentiment or score). Lighter colors (yellow/green) typically indicate higher values, while darker colors (blue) indicate lower values.
From the visualization:

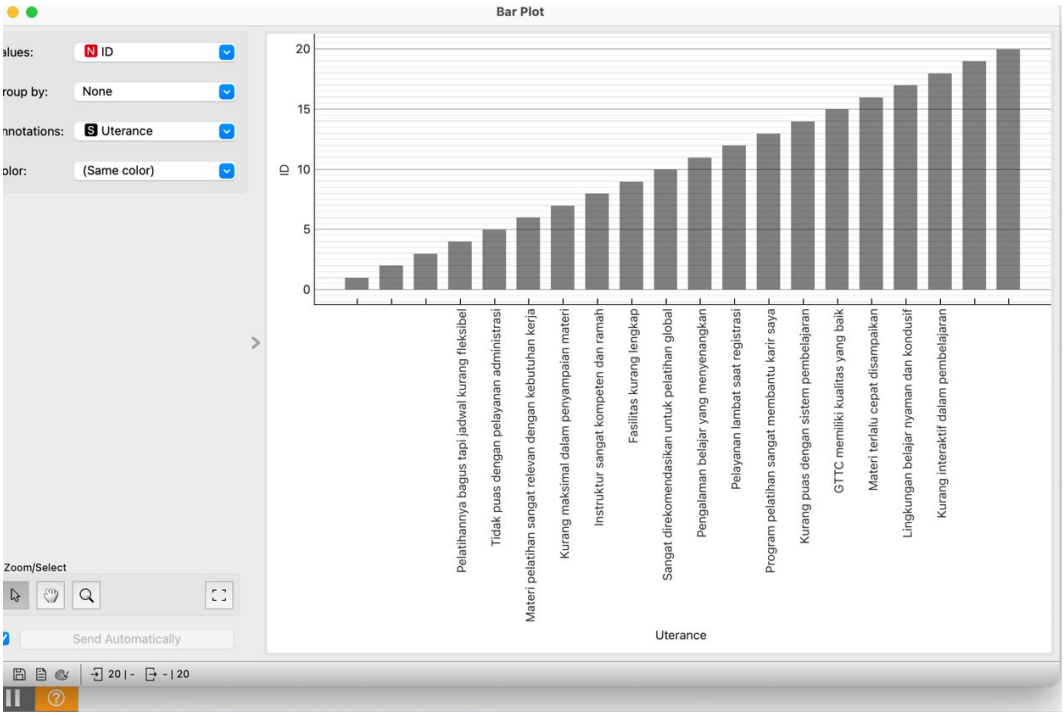
- Statements such as “Very satisfied with the training results” or “GTTC has good quality” tend to have high (positive) scores.
- On the other hand, statements such as “Not optimal in delivering material” or “Slow service during registration” show lower scores (indications of dissatisfaction).

This heat map helps identify areas for improvement (e.g., administrative services or speed of material delivery) as well as aspects that are already working well .

g. *Plot*

A bar plot is a visualization of the results of text data processing using Orange Data Mining, presenting respondents' statements or comments in the form of a bar chart. Each bar represents a single comment arranged sequentially based on its data identity (ID). The length of the bar does not reflect the number of words or sentiment value, but rather indicates the position of each statement within the overall dataset. This visualization helps understand the arrangement of respondents' comments in a more systematic, structured, and easily analyzed manner.

Figure 10. Bar Plot



Source: Orange Data Mining Software

A bar chart displays the number or frequency of each statement (utterance). The horizontal axis lists the statements, while the vertical axis shows the number of occurrences or a specific level.

From the graph:

- Statements with the highest bars indicate the most frequently occurring issues or opinions, such as “Lack of interactivity in learning” or “Comfortable and conducive learning environment”.
- Statements with lower stems indicate issues that are mentioned relatively infrequently.

This visualization is useful for:

- Identify priority issues or strengths based on frequency.
- Determine the aspects that receive the most attention from respondents.

CONCLUSION

Based on the results of research that has been conducted regarding sentiment analysis of the reputation of the Global Training and Test Center (GTTC) Sukabumi using a text mining and sentiment analysis approach based on Orange Data Mining, several things can be concluded as follows:

Sentiment analysis results indicate that stakeholder perceptions of the Sukabumi GTTC tend to be dominated by positive sentiment. This is evident in the numerous comments containing terms such as "satisfied," "professional," and "useful," reflecting participants' level of satisfaction with the quality of service provided. This positive sentiment indicates that the GTTC has succeeded in building a positive public image. Despite this, negative and neutral sentiments were still found, indicating several aspects that need improvement. Some of the issues that emerged included the lack of interactivity in learning methods, slow administrative services, and suboptimal material delivery. This demonstrates the need for ongoing evaluation to improve overall service quality. Topic modeling results identified that the main factors influencing GTTC's reputation include training quality, facilities, service, and instructor competence. The dominant words in the word cloud reinforce the finding that participants' experiences during the training are a key factor in shaping their perceptions of the institution. The use of text mining and sentiment analysis approaches has proven effective in transforming unstructured text data into strategically valuable information. This method not only identifies sentiment trends but also provides deeper insights into the factors influencing stakeholder opinions.

Practically, the results of this study can serve as a basis for GTTC management in formulating strategies to improve service quality. Improvements can be focused on administrative services, more interactive learning methods, and enhancing the quality of material delivery. Meanwhile, aspects that have been assessed as positive need to be maintained and continuously developed to strengthen the institution's reputation on an ongoing basis. Thus, this study confirms that the use of text mining-based data analysis technology has an important role in supporting more objective, data-based, and adaptive decision-making to stakeholder needs in the digital era.

REFERENCES

Anwar, M., Hidayat, T., & Prasetyo, A. (2024). Integration of sentiment analysis and topic modeling for organizational reputation assessment. Journal of Data Science and Analytics, 12(2), 115–128.

Hidayat, R., Nugroho, A., & Saputra, D. (2024). Sentiment analysis in education services: A data-driven approach to user satisfaction. International Journal of Educational Technology, 9(1), 45–58.

Kurniawan, D., Putra, R., & Wijaya, S. (2023). Topic modeling using Latent Dirichlet Allocation (LDA) for large-scale text analysis. Journal of Information Systems, 18(2), 89–102.

Prasetyo, B., Santoso, H., & Lestari, D. (2022). Visual data mining using Orange software for text analytics. Indonesian Journal of Computer Science, 7(1), 23–34.

Pratama, A., Rahman, F., & Sari, N. (2023). Text mining techniques for extracting insights from unstructured data. Journal of Big Data Research, 5(2), 67–80.

Putri, D., & Santoso, B. (2022). The role of online reviews in shaping organizational reputation. Journal of Business and Management Studies, 10(3), 210–220.

Rachmawati, L., & Nugroho, Y. (2021). The impact of text preprocessing on sentiment analysis accuracy. Journal of Artificial Intelligence Research, 6(1), 33–41.

Rahman, A., & Putra, M. (2021). Implementation of Latent Dirichlet Allocation in document clustering. International Journal of Computer Applications, 174(5), 12–18.

Rani, P., & Arora, V. (2016). Text preprocessing techniques in text mining. International Journal of Computer Applications, 133(6), 7–10.

Sari, M., & Wijaya, A. (2022). Sentiment analysis using VADER for social media data. Journal of Informatics and Data Science, 4(2), 55–63.

Sari, N., Pratama, A., & Hidayat, R. (2023). Application of sentiment analysis for measuring customer satisfaction. Journal of Information Technology, 8(1), 14–25.

Wijaya, R., & Putra, D. (2022). Lexicon-based sentiment analysis for Indonesian text. Procedia Computer Science, 179, 345–352.