Journal of Information System Exploration and Research

https://joiser.shmpublisher.com

p-ISSN 2964-1160 | e-ISSN 2963-6361

Student Sentiment Analysis on the Use of Generative Artificial Intelligence Platforms to Support Academic Activities

Riva Ariefta[1*], Heny Pratiwi[2], Ahmad Fahrijal Pukeng[3]

[1,2,3]*Department of Information Systems, STMIK Widya Cipta Dharma, Indonesia*

DOI: https://doi.org/10.52465/joiser.v4i3.102

Received 10 July 2026; Accepted 04 August 2026; Available online 07 August 2026

Article Info

Keywords:

Sentiment analysis;
Generative artificial intelligence;
Naïve bayes;
TF-IDF;
Text mining

Abstract

The development of Generative Artificial Intelligence (Generative AI) has significantly supported students' academic activities, including information retrieval, assignment completion, and learning. This study aims to analyze student sentiment toward the use of Generative AI platforms in academic contexts using the Naïve Bayes algorithm. Unlike previous studies that primarily examined social media data or focused on a single Generative AI platform, this research utilizes primary textual responses collected directly from students regarding their experiences with multiple Generative AI platforms, providing a broader understanding of student perceptions in higher education. A quantitative approach was employed using questionnaire data from 102 students. After data screening, two incomplete responses were excluded, resulting in 100 valid responses for sentiment analysis. The data underwent text preprocessing, including case folding, cleaning, tokenization, stopword removal, and stemming, followed by TF-IDF weighting before classification. The results indicate that 73% of responses expressed positive sentiment, 22% were neutral, and 5% were negative. Model evaluation achieved an accuracy of 75%, precision of 56%, recall of 75%, and an F1-score of 64%. Overall, the findings demonstrate that students generally perceive Generative AI platforms positively and consider them beneficial in supporting academic activities, highlighting their growing role in higher education learning environments.

1. Introduction

The development of Generative Artificial Intelligence (Generative AI) has given rise to various platforms used in higher education, including ChatGPT, Gemini, and Claude. These three platforms are used to help students understand course material, find references, organize assignments, and improve the effectiveness of the learning process through natural language-based interactions. [1], [2], [3], [4],

[*] ***Corresponding Author:***

Riva Ariefta,
Department of Information Systems,
STMIK Widya Cipta Dharma Samarinda,
Samarinda, East Kalimantan, Indonesia.
Email: 2341037@wicida.ac.id

[5]. However, the quality of the information produced must be continuously evaluated because there is a possibility of inaccurate information and various problems related to academic ethics [6].

Various studies show that Generative AI can improve learning effectiveness and student productivity [7]. However, user perceptions of this technology vary. Some students find Generative AI very helpful in the learning process, while others are concerned about the accuracy of the information, the reliance on the technology, and its impact on critical thinking skills [8]. Furthermore, based on a preliminary interview with the Head of the Information Systems Study Program at STMIK Widya Cipta Dharma Samarinda, the aspects considered most important in choosing a Generative AI platform include answer accuracy, ease of use, response speed, feature availability, and cost or accessibility. These findings suggest that student perceptions of the use of Generative AI platforms require further study to determine their sentiments toward this technology in supporting academic activities. Therefore, research is needed that can identify student views on the use of Generative AI platforms in academic activities [9].

Sentiment analysis is a technique in Natural Language Processing (NLP) that is used to identify user opinions based on positive, negative, or neutral categories [10]. One of the algorithms that is widely applied in sentiment analysis is Naïve Bayes, because it has a simple computational process and is able to provide good classification performance on text data [11]. Although numerous studies have been conducted on sentiment analysis using the Naïve Bayes algorithm, most have focused on analyzing app reviews, social media reviews, or evaluating a single Generative AI platform. Research analyzing student opinions on the use of various Generative AI platforms using primary data in the form of questionnaires in the context of academic activities is still limited. Therefore, this study aims to fill this gap by analyzing student sentiment using the Naïve Bayes algorithm [12].

Based on these problems, student sentiment analysis towards the use of the Generative Artificial Intelligence platform in supporting academic activities was carried out using the Naïve Bayes algorithm [13]. The research data was obtained through a questionnaire containing student opinions, then processed through text preprocessing, TF-IDF weighting, and classification using Naive Bayes. The research results are expected to provide an overview of students' views on the use of Generative AI in academic activities and provide input for universities in optimizing the application of AI technology in the learning process [8].

2. Literature Review

This section presents the theoretical foundation that underpins this study. It discusses the concepts of Generative Artificial Intelligence in education, sentiment analysis, text preprocessing and TF-IDF weighting, the Naïve Bayes algorithm, and previous related studies. These theories provide the conceptual basis for analyzing student sentiment toward the use of Generative AI platforms in supporting academic activities.

2.1 Generative artificial intelligence in education

Generative Artificial Intelligence (Generative AI) is a branch of artificial intelligence that is capable of generating various types of content, such as text, images, and program code based on commands (prompts) given by the user [14]. Since the introduction of ChatGPT in 2022, the development of Generative AI has been increasingly rapid and has been followed by the presence of other platforms such as Gemini and Claude [15]. In higher education settings, these three platforms are used by students to help them understand course materials, find references, prepare assignments, and support the completion of various academic activities more effectively and efficiently [16], [17].

Despite its numerous benefits, the use of Generative AI also presents a number of challenges, such as the potential for inaccurate information delivery, user dependence on technology, and the emergence of academic ethics issues, including plagiarism [18]. Therefore, the information generated by Generative AI still needs to be verified using trusted sources [8]. Evaluation of student perceptions and sentiments is also necessary to determine how this technology is utilized to support academic activities optimally and responsibly [19].

2.2 Sentiment analysis

Sentiment analysis is a Natural Language Processing (NLP) technique used to identify and classify user opinions into positive, negative, or neutral categories [20]. In education, sentiment analysis helps understand students' perceptions of the learning process, academic services, and the implementation of new technologies. Sentiment analysis results can inform evaluations of educational institutions' user experiences and support better decision-making.

2.3 Text preprocessing and TF-IDF

Text data obtained from questionnaires generally still contains irrelevant words, spelling variations, and characters that can affect the analysis process. Therefore, a text pre-processing stage is required to clean and standardize the text data. The stages carried out include case folding, cleaning, tokenization, stopword removal, and stemming. After the text pre-processing process is complete, the text data is converted into a numeric representation using the Term Frequency–Inverse Document Frequency (TF-IDF) method. The TF-IDF method assigns weights to each word based on its frequency of occurrence in a document and its occurrence rate across documents, so that words with a higher level of importance will receive greater weight. The resulting numeric representation is then used as an input feature in the sentiment classification process using the Naïve Bayes algorithm [16], [17], [21].

2.4 Naïve bayes algorithm

Naïve Bayes is a classification algorithm based on Bayes' Theorem which is widely used in sentiment analysis research because it is simple, efficient, and has good performance in classifying text data [22]. This algorithm calculates the probability that a document belongs to a particular sentiment class based on the distribution of words appearing in the training data. Several studies have shown that the combination of text preprocessing, TF-IDF weighting, and the Naive Bayes algorithm can produce good classification performance on user opinion data and generative AI application reviews.

2.5 Previous research

Research on sentiment analysis has been widely applied to identify user opinions on various digital services, applications, and artificial intelligence-based technologies. One widely used algorithm is Naïve Bayes because it has a simple computational process, is easy to implement, and is capable of delivering good performance in text data classification. The combination of Term Frequency–Inverse Document Frequency (TF-IDF) and Naïve Bayes has also proven effective in improving sentiment classification results in various studies [23], [24]**.**

However, research on sentiment toward the use of Generative Artificial Intelligence (AI) is still dominated by evaluations of specific platforms or data from social media. Research using primary data, such as student opinions, regarding the use of various Generative AI platforms to support academic activities, is still relatively limited [25].

A preliminary interview with the Head of the Information Systems Study Program at STMIK Widya Cipta Dharma Samarinda revealed five key aspects in selecting a Generative AI platform: answer accuracy, ease of use, response speed, feature availability, and cost or accessibility. These aspects indicate that the assessment of a Generative AI platform is influenced not only by information quality but also by the user experience in supporting academic activities.

Based on these conditions, this study offers a different approach from previous research by utilizing primary data obtained directly from students in the form of opinions regarding the use of various Generative AI platforms to support academic activities. The opinion data was then processed through text preprocessing, TF-IDF weighting, and sentiment classification using the Naïve Bayes algorithm. The results are expected to provide a general overview of student sentiment tendencies towards the use of Generative AI platforms and provide input for universities in formulating policies for the effective and responsible use of AI technology.

3. Method

This section describes the research methodology employed in this study. It explains the research approach, research objects and respondents, instrument development, data collection procedures, and data analysis techniques used to analyze student sentiment toward the use of Generative Artificial Intelligence platforms.

3.1 Research approach

This study employed a quantitative approach with sentiment analysis to identify students' opinions regarding the use of Generative Artificial Intelligence (Generative AI) platforms in supporting academic activities. The Naïve Bayes algorithm was selected because it has a simple computational process, is efficient for text classification, and has demonstrated reliable performance in various sentiment analysis studies involving textual data. In addition, Naïve Bayes performs effectively when combined with text preprocessing and TF-IDF weighting, making it one of the most widely used classification methods in text mining research [11], [21], [22], [23].

3.2 Research objects and respondents

Objek The object of this research is the use of Generative Artificial Intelligence (Generative AI) platforms, such as ChatGPT, Gemini, and Claude, to support student academic activities. The respondents were 102 Information Systems students at STMIK Widya Cipta Dharma Samarinda who had used one or more Generative AI platforms to support their learning.

3.3 Preparation of research instruments

Prior to developing the research instrument, a preliminary interview was conducted with the Head of the Information Systems Study Program at STMIK Widya Cipta Dharma Samarinda to identify aspects to consider when using the Generative AI platform. Based on the interview results, five key indicators were identified: answer accuracy, ease of use, response speed, feature availability, and accessibility. These five indicators formed the basis for developing the questionnaire.

The research instrument consisted of two types of data: quantitative and textual data. Quantitative data were obtained through ten statements structured around the five indicators from the preliminary interview and measured using a five-point Likert scale. Before being used in the research, the instrument's validity was tested using the Pearson Product Moment correlation to determine the ability of each statement item to measure the research variables. Furthermore, reliability was tested using the Cronbach's Alpha coefficient to measure instrument consistency. The instrument was declared valid if the calculated r value was greater than the table r value at the 5% significance level. The instrument was considered reliable if the Cronbach's Alpha value was greater than the table r value at the 5% significance level.0.

3.4 Data collection

Data collection was conducted online using Google Forms. A total of 102 questionnaires were collected from Information Systems students. After the data screening process, two responses were excluded because the open-ended opinion fields were incomplete and therefore unsuitable for text mining analysis. Consequently, 100 valid textual responses were used in the sentiment analysis process. In addition to identity questions and Likert-type statements, the questionnaire also included open-ended questions asking respondents to describe their experiences using the Generative AI platform to support academic activities. The answers to these open-ended questions served as the primary data for the sentiment analysis process.

3.5 Data analysis

The obtained opinion data was then assigned sentiment labels into three categories: positive, neutral, and negative, based on the content and meaning of respondents' responses. The labeling process was performed manually by researchers as reference data for the classification process. Next, the text data underwent text preprocessing, which included case folding, cleaning, tokenization, stopword removal, and stemming. After the preprocessing process was complete, word weighting was performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method.

The TF-IDF weighting results were then divided into training data (80%) and testing data (20%). The 80:20 split was selected because it is one of the most commonly used data partitioning strategies in supervised machine learning, providing sufficient data for model training while maintaining an adequate proportion of unseen data for performance evaluation. This approach has been widely adopted in text classification and sentiment analysis studies to achieve reliable model evaluation [21], [23]. The training data were used to build the Naïve Bayes classification model, whereas the testing data were used to evaluate the model performance.

Model performance was evaluated using a confusion matrix by calculating accuracy, precision, recall, and F1-score. These evaluation metrics are widely used to assess classification performance because they provide comprehensive information regarding prediction correctness, class identification capability, and the balance between precision and recall, particularly in text classification and sentiment analysis tasks [11], [21], [24]. The evaluation results were used to determine the ability of the Naïve Bayes algorithm to classify student sentiment toward the use of the Generative Artificial Intelligence platform. The overall research stages are shown in Figure 1.

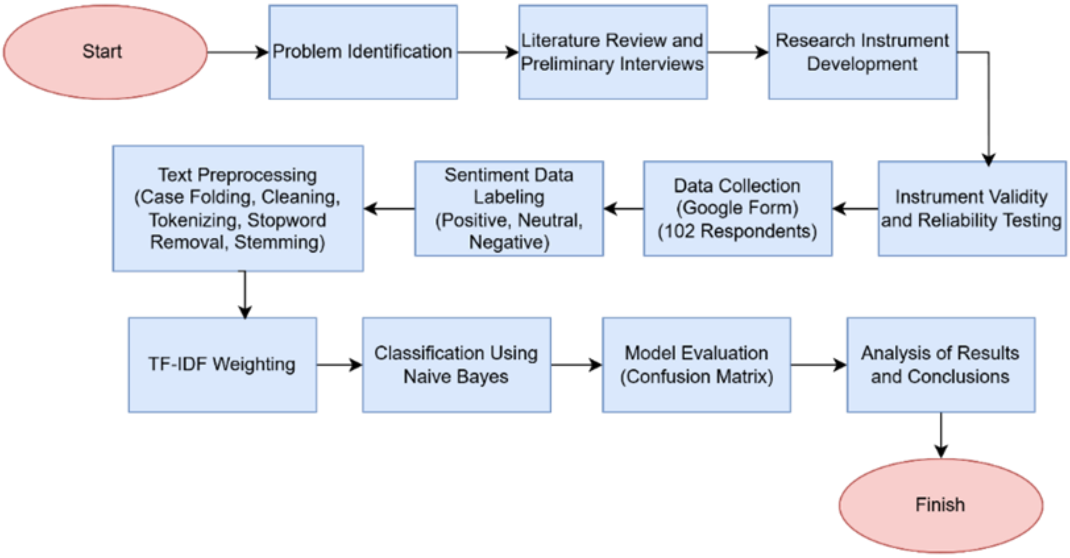

Figure 1. Research flow

Table 1. Research instrument indicators

Variable	Indicator	Code Number	Number of Grains
Use of Generative Artificial Intelligence Platform	Answer Accuracy	P1, P2	2
Use of Generative Artificial Intelligence Platform	Ease of Use	P3, P4	2
Use of Generative Artificial Intelligence Platform	Response Speed	P5, P6	2
Use of Generative Artificial Intelligence Platform	Feature Availability	P7, P8	2
Use of Generative Artificial Intelligence Platform	Accessibility	P9, P10	2

4. Results and Discussion

This section presents the results of the study based on the collected questionnaire data. The findings are described using descriptive statistical analysis to summarize the respondents' characteristics and questionnaire responses. In addition, text mining and sentiment analysis are conducted to explore respondents' perceptions of the use of Generative Artificial Intelligence (GenAI) in academic activities.

4.1 Respondent characteristics

This study initially involved 102 Information Systems students at STMIK Widya Cipta Dharma Samarinda who had used one or more Generative Artificial Intelligence platforms in their academic activities. Respondent characteristics were analyzed using data from all 102 questionnaires. After the data screening process, two incomplete open-ended responses were excluded from the sentiment analysis, resulting in 100 valid textual responses for text mining.

Percentage Formula shown in Equation (1):

$$P = \frac{f}{N} \times 100\% \quad (1)$$

P = Percentage of respondents (%)
f = Frequency or number of respondents in each category
N = Total number of respondents
100% = Conversion to a percentage

Table 2. Respondent characteristics based on gender

Gender	Amount	Percentage (%)
Male	55	53,92
Female	47	46,08
Total	102	100

Based on Table 2, it is known that of the 102 respondents, 55 respondents (53.92%) were male, while 47 respondents (46.08%) were female. These results indicate that the number of male respondents slightly outnumbered female respondents in this study.

Table 3. Characteristics of respondents' most frequently used generative AI platforms

Generative AI Platform	Frequency	Percentage (%)
ChatGPT	63	61,76
Gemini	25	24,51
Claude	10	9,80
Lainnya	4	3,92

Based on Table 3, the majority of respondents used ChatGPT as the Generative AI platform most frequently used in academic activities, namely 63 respondents (61.76%), followed by Google Gemini with 25 respondents (24.51%), and Claude with 10 respondents (9.80%). A total of 4 respondents (3.92%) chose the Other category, which contains free answers such as using multiple platforms simultaneously or answers that do not specifically mention the platform name.

Table 4. Respondent characteristics based on frequency of use

Frequency of use	Frequency	Percentage (%)
Every day	28	27,45
Several times a week	60	58,82
Several times a month	6	5,88
Seldom	8	7,84

Based on Table 4, the majority of respondents (60 respondents) use the Generative Artificial Intelligence platform several times a week. Furthermore, 28 respondents (27.45%) use the platform daily, 8 respondents (7.84%) use it infrequently, and 6 respondents (5.88%) use the platform several times a month. These results indicate that the majority of students have routinely utilized Generative AI to support their academic activities.

4.2 Validity and reliability test results

Prior to conducting sentiment analysis, the research instrument was first tested for validity and reliability to ensure that each statement accurately and consistently measured the researched aspect. Validity was tested using the Pearson Product Moment method, while reliability was tested using the Cronbach's Alpha coefficient. Testing was conducted on 10 statements using data from 102 respondents.

Table 5. Instrument validity test results

No	code	r count	r table	Information
1	P1	0,548	0,195	Valid
2	P2	0,467	0,195	Valid
3	P3	0,715	0,195	Valid
4	P4	0,773	0,195	Valid
5	P5	0,755	0,195	Valid
6	P6	0,788	0,195	Valid
7	P7	0,726	0,195	Valid
8	P8	0,722	0,195	Valid
9	P9	0,731	0,195	Valid
10	P10	0,597	0,195	Valid

Based on Table 5, all statement items have calculated r values greater than the table r value of 0.195 at a 5% significance level. The calculated r values are in the range of 0.467–0.788, so all statement items are declared valid and suitable for use as research instruments.

Table 6. Instrument reliability test results

Cronbach's Alpha	Number of Items	Informations
0,913	10	Reliabel

Based on Table 6, the Cronbach's Alpha value is 0.913. This value is greater than the minimum threshold of 0.70, thus indicating that the research instrument is reliable. These results indicate that all statement items have an excellent level of internal consistency and are suitable for use in research data collection.

4.3 Result of student perception analysis

Student perception analysis was conducted on ten statements structured around five indicators: answer accuracy, ease of use, response speed, feature availability, and accessibility. Each statement was scored using a five-point Likert scale: Strongly Disagree (1), Disagree (2), Neutral (3), Agree (4), and Strongly Agree (5). This analysis aimed to determine student perceptions regarding the use of Generative Artificial Intelligence platforms to support academic activities.

The score for each indicator was obtained by calculating the average (mean) of the respondents' responses. The average score was calculated using the following shown in Equation (2):

$$\bar{x} = \frac{\sum X}{N} \tag{2}$$

Description:
$\bar{x}$ = avarage value (mean)
$\sum X$ = total score of all respondents' answers
N = number of respondents

The average value obtained is then interpreted into assessment categories based on Likert scale intervals. Category intervals are calculated using the formula shown in Equation (3) and (4):

$$I = \frac{S_t - S_r}{K} \tag{3}$$

$$\bar{x} = \frac{\sum X}{N} \tag{4}$$

Description:
I = Width of the category interval
St = Highest score (5)
Sr = Lowest score (1)
K = Number of categories (5)

Table 7. Mean value interpretation category

Interval Mean	Category
4,21 – 5,00	Very Good
3,41 – 4,20	Good
2,61 – 3,40	Fairly Good
1,81 – 2,60	Poor
1,00 – 1,80	Very Poor

As a basis for interpreting the results of the student perception analysis, the average scores were classified into five assessment categories based on Likert scale intervals. These categories are presented in Table 7 and used to determine respondents' level of perception regarding each indicator studied.

Table 8. Results of student perception analysis based on indicators

Indicator	Question	Mean	Category
Answer Accuracy	P1 and P2	3,65	Good
Ease of Use	P3 and P4	4,25	Very Good
Response Speed	P5 and P6	4,00	Good
Feature Availability	P7 and P8	3,85	Good
Accessibility	P9 and P10	3,91	Good

The analysis results in Table 8 show that the ease of use indicator obtained the highest average score of 4.25, categorized as Very Good. This indicates that students consider the Generative Artificial Intelligence platform easy to use to support academic activities. On the other hand, the answer accuracy indicator obtained the lowest average score of 3.65, but still falls into the Good category. Overall, all indicators are in the Good to Very Good category, thus it can be concluded that students have a positive perception of the use of the Generative Artificial Intelligence platform in academic activities.

4.4 Sentiment analysis results

Although 102 questionnaires were collected, only 100 textual responses were eligible for sentiment analysis after the data screening process because two respondents provided incomplete open-ended answers. Sentiment analysis was conducted on these 100 valid responses regarding students' experiences using Generative Artificial Intelligence platforms to support academic activities. Each response was manually labeled into positive, neutral, and negative sentiment categories before being used as the ground truth for the Naïve Bayes classification process.

Table 9. Distribution of sentiment data

Sentiment	Amount	Percentage (%)
Positive	73	73,00
Neutral	22	22,00
Negative	5	5,00
Total	100	100,00

The sentiment analysis results in Table 9 show that of the 100 respondents' opinion data, 73 (73%) were positive, 22 (22%) were neutral, and 5 (5%) were negative. The dominance of positive sentiment indicates that the majority of students responded favorably to the use of Generative Artificial Intelligence to support academic activities. Meanwhile, negative sentiment indicates that a small number of respondents still believe Generative AI has limitations, such as inconsistent response accuracy or information that still needs to be verified.

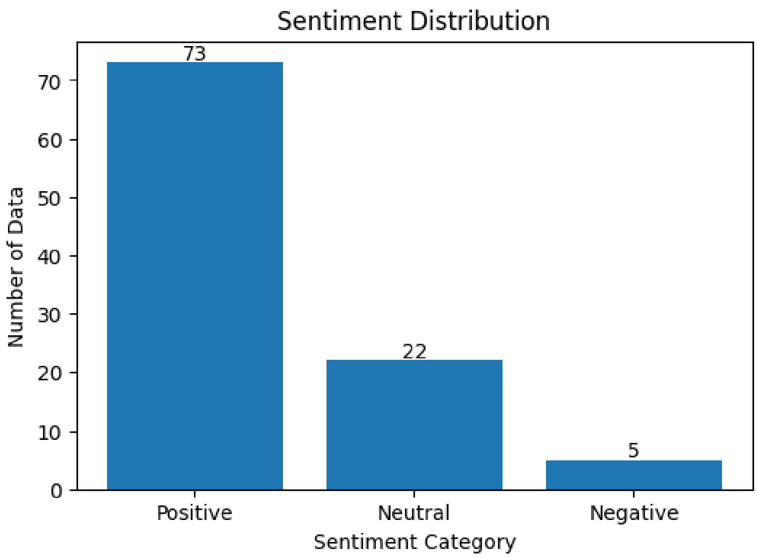

Figure 2. Sentiment distribution diagram

Based on the results of labeling 100 respondents' comments, three sentiment categories were obtained: positive sentiment (73 data points), neutral sentiment (22 data points), and negative sentiment (5 data points). The dominance of positive sentiment indicates that the majority of respondents responded positively to the use of the Generative Artificial Intelligence platform in supporting academic activities. Meanwhile, the presence of negative sentiment indicates that a small number of respondents still expressed dissatisfaction or criticism, such as regarding the accuracy of answers, limited features, or the need to verify the information generated. The distribution of sentiment labeling results is presented in Figure 2.

4.5 Text preprocessing results

Before the classification process using the Naïve Bayes algorithm, the respondent opinion data first went through a text preprocessing stage. This stage aims to clean and standardize the text data to obtain a more structured word representation and reduce words that have no significance in the classification process. In this study, the text preprocessing stages included case folding, cleaning, tokenizing, stopword removal, and stemming. Examples of the results of each text preprocessing stage are presented in Table 10.

Table 10. Example of text preprocessing results

Stages	Results
Original Sentence	It is certainly very helpful because it is more time-efficient.
Case Folding	it is certainly very helpful because it is more time-efficient
Cleaning	it is certainly very helpful because it is more time efficient
Tokenizing	it, is, certainly, very, helpful, because, it, is, more, time, efficient
Stopword Removal	certainly very helpful more time efficient
Stemming	certainly helpful efficient time

Table 10 shows the changes in text data at each stage of text preprocessing. The case folding stage converts all letters to lowercase to ensure a uniform writing format. Next, cleaning removes unnecessary characters, while tokenizing separates sentences into tokens or words. The stopword removal stage removes common words that do not contribute significantly to the classification process. Finally, stemming converts each word to its base form, making the text data more structured and ready for use in the TF-IDF weighting process and classification using the Naïve Bayes algorithm.

4.6 TF-IDF weighting results

After going through the text preprocessing stage, the text data is converted into a numerical representation using the Term Frequency–Inverse Document Frequency (TF-IDF) method. This method is used to assign weights to each word based on its occurrence rate in the document and its frequency across all documents. The TF-IDF weighting results are then used as input for the classification process using the Naïve Bayes algorithm.

Table 11. Top 10 terms with the highest TF-IDF weights

No	Words	TF-IDF Weight
1	very	11.446767
2	help	11.206766
3	AI	6.579127
4	good	6.340715
5	quite	5.460938
6	easy	5.255500
7	good	5.094711
8	assignment	4.518755
9	use	4.405104
10	academic	4.200312

Based on Table 11, the word "very" has the highest TF-IDF weight, which is 11.446767, followed by the word "help" with a weight of 11.206766. High weight values indicate that these two words frequently appear in respondents' opinion data and have an important role in representing the content of the document. In addition, the words "ai", "good", "enough", "easy", "good", "task", "use", and "academic" also have relatively high TF-IDF weights. This indicates that these words are important features in describing students' perceptions of the use of the Generative Artificial Intelligence platform. Furthermore, the TF-IDF weighted results are used as input features in the sentiment classification process using the Naïve Bayes algorithm.

4.7 Naïve bayes classification results

After undergoing text preprocessing and weighting using the Term Frequency–Inverse Document Frequency (TF-IDF) method, the data was then classified using the Naïve Bayes algorithm. The classification process was carried out by dividing the data into 80% training data and 20% testing data. The training data was used to build the classification model, while the testing data was used to evaluate the model's ability to classify student sentiment toward the use of the Generative Artificial Intelligence platform.

The model evaluation results showed that the Naïve Bayes algorithm achieved an accuracy of 75%, meaning the model was able to correctly classify 75% of the test data. Furthermore, the weighted precision was 0.56, the weighted recall was 0.75, and the weighted F1-score was 0.64. The model evaluation results are presented in Table 12.

Table 12. Naïve bayes model evaluation results

Metrik	Value
Accuracy	75%
Precision	56%
Recall	75%
F1-Score	64%

In Table 12, the Naïve Bayes model achieved an accuracy of 75%, a precision of 56%, a recall of 75%, and an F1-score of 64%. These results indicate that the model was able to classify most of the test data well. However, based on the classification report, the model's performance in the negative and neutral sentiment classes was still low because the amount of data in these two classes was relatively small compared to the positive class. This condition caused the model to be more likely to predict data in the positive class.

To understand the classification results in more detail, an evaluation was conducted using the confusion matrix presented in Table 13.

Table 13. Confusion matrix

Prediction	Negative	Neutral	Positive
negative	0	0	1
positive	0	0	4
neutral	0	0	15

The confusion matrix results in Table 13 show that the model successfully classified 15 pieces of positive sentiment data correctly. However, the model was unable to identify data in the negative and neutral sentiment classes, resulting in one negative sentiment data item and four neutral sentiment data items being predicted as positive. This indicates that the model still tends to classify data into the positive sentiment class. This is likely influenced by an imbalance in data distribution, where positive sentiment data outnumbers neutral and negative sentiment data.

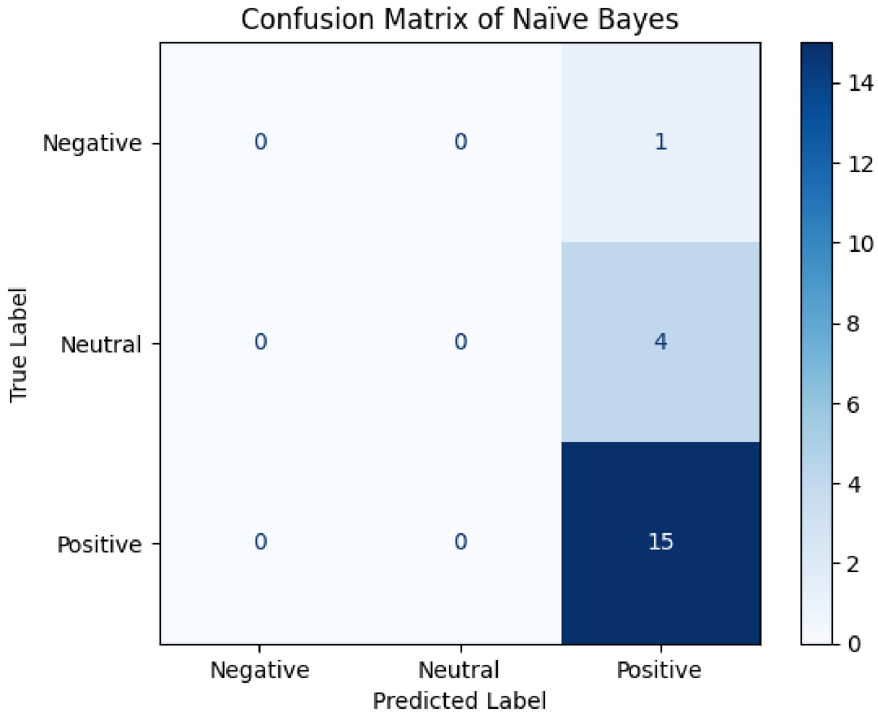

Figure 3. Confusion matrix

Figure 3 shows the confusion matrix results from the Naïve Bayes algorithm on the test data. The model successfully classified 15 positive sentiment data sets correctly. However, 1 negative sentiment data set and 4 neutral sentiment data sets were still predicted as positive. These results indicate that the model performs better in recognizing positive sentiment than negative and neutral sentiment. This is likely due to the imbalanced data distribution, where positive sentiment data sets outnumber the other two classes.

4.8 Discussion

The study results show that the majority of students have a positive perception of the use of Generative Artificial Intelligence (Generative AI) platforms to support academic activities. These findings indicate that Generative AI helps students find information, understand material, and complete academic assignments. These results align with research by Almassaad et al.[25] which states that Generative AI provides benefits in improving the learning process in higher education.

Sentiment analysis shows that 73% of the data falls into the positive category, 22% into the neutral category, and 5% into the negative category. The predominance of positive sentiment indicates that most respondents had a positive experience using Generative AI. Meanwhile, the presence of negative sentiment indicates that a small number of students still gave less positive assessments, such as those related to inconsistent answer accuracy or the need to verify the generated information. This finding is also supported by Chan and Hu [8] who reported that students tended to have positive perceptions about the use of Generative AI in the learning process.

In the TF-IDF weighting process, words such as "very", "helpful", "AI", "good" and "easy" received the highest weights. This indicates that usability and ease of use were the aspects most frequently mentioned by respondents. According to Gilbert et al.[26] TF-IDF weighting is able to produce effective feature representation, thus supporting the sentiment classification process using the Naïve Bayes algorithm.

Tests using the Naïve Bayes algorithm yielded an accuracy score of 75%, indicating that the model was able to correctly classify most of the test data. However, the confusion matrix results indicate that the model still has difficulty distinguishing negative and neutral sentiments, as both classes tend to be predicted as positive. This condition is suspected to be influenced by the imbalanced data distribution, where positive sentiment data outnumbers neutral and negative sentiment data. Mulyani and Novita [27] also explains that the imbalance in the amount of data can affect the performance of the Naïve Bayes algorithm in the classification process.

5. Conclusion

This study successfully applied sentiment analysis using the Naïve Bayes algorithm to analyze student perceptions of the use of a Generative Artificial Intelligence (Generative AI) platform to support academic activities. The analysis results showed that the majority of students expressed positive sentiments toward the use of Generative AI, indicating that the platform is considered capable of assisting the learning process, information retrieval, and completion of academic assignments.

The application of text preprocessing, TF-IDF weighting, and classification using the Naïve Bayes algorithm resulted in a model with 75% accuracy. These results indicate that the Naïve Bayes algorithm was able to correctly classify most of the sentiment data. However, the confusion matrix results indicated that the model still had difficulty distinguishing negative from neutral sentiment due to the imbalanced data distribution, where positive sentiment data was more dominant than other classes.

Overall, this study demonstrates that the combination of the TF-IDF method and the Naïve Bayes algorithm can be used to analyze student sentiment toward the use of Generative AI to support academic activities. It is recommended to conduct further research using larger amounts of data, more balanced class distributions, and comparing the Naïve Bayes algorithm with other classification methods to improve model performance.

Use of Artificial Intelligence

The authors used ChatGPT (OpenAI) solely for grammar and language refinement during the preparation of this manuscript. The use of artificial intelligence was limited to language editing assistance. All research design, data collection, data analysis, interpretation of results, and scientific conclusions were performed and verified entirely by the authors.

Credit Authorship Contribution Statement

Riva Ariefta: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing Original Draft. **Heny Pratiwi**: Supervision, Validation, Writing, Review & Editing. **Ahmad Fahrijal Puken**: Supervision, Validation, Writing, Review & Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

References

[1] K. Nikolopoulou, "Generative Artificial Intelligence in Higher Education: Exploring Ways of Harnessing Pedagogical Practices with the Assistance of ChatGPT," *Int. J. Chang. Educ.*, vol. 1, no. 2, pp. 103–111, 2024, doi: 10.47852/bonviewijce42022489.

[2] L. Isiaku, A. S. Muhammad, H. I. Kefas, and F. C. Ukaegbu, "Enhancing technological sustainability in academia: leveraging ChatGPT for teaching, learning and evaluation," *Qual. Educ. All*, vol. 1, no. 1, pp. 385–416, 2024, doi: 10.1108/qea-07-2024-0055.

[3] D. Baidoo-Anu, D. Asamoah, I. Amoako, and I. Mahama, "Exploring student perspectives on generative artificial intelligence in higher education learning," *Discov. Educ.*, vol. 3, no. 1, 2024, doi: 10.1007/s44217-024-00173-z.

[4] Q. Liu, A. Hu, T. Gladman, and S. Gallagher, "Eight Months into Reality: A Scoping Review of the Application of ChatGPT in Higher Education Teaching and Learning," *Innov. High. Educ.*, vol. 50, no. 5, pp. 1677–1700, 2025, doi: 10.1007/s10755-025-09790-4.

[5] H. Pratiwi, "Implementation of Cloud-Based Management Information Systems for Optimizing Educational Administration in Higher Education," *Inf. Syst. Anal. Des. Dev.*, vol. 1, no. 2, pp. 67–77, 2026, doi: 10.66472/isadd.v1i2.56.

[6] G. A. Wibowo, A. Rahman, and M. Anis, “The Impact of ChatGPT Use on the Quality of Academic Support for Students,” *Technol. Soc. Perspect.*, vol. 1, no. 3, pp. 132–138, 2023, doi: 10.61100/tacit.v1i3.69.
[7] A. Bozkurt, “GenAI et al.: Cocreation, Authorship, Ownership, Academic Ethics and Integrity in a Time of Generative AI,” *Open Prax.*, vol. 16, no. 1, pp. 1–10, 2024, doi: 10.55982/openpraxis.16.1.654.
[8] C. K. Y. Chan and W. Hu, “Students’ voices on generative {AI}: Perceptions, benefits, and challenges in higher education,” *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 43, 2023, doi: 10.1186/s41239-023-00411-8.
[9] T. K. F. Chiu, “Future research recommendations for transforming higher education with generative AI,” *Comput. Educ. Artif. Intell.*, vol. 6, p. 100197, 2023, doi: 10.1016/j.caeai.2023.100197.
[10] M. S. U. Miah, M. M. Kabir, T. Bin Sarwar, M. Safran, S. Alfarhood, and M. F. Mridha, “A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–18, 2024, doi: 10.1038/s41598-024-60210-7.
[11] M. Romano, G. Contu, F. Mola, and C. Conversano, “Threshold-based Naïve Bayes classifier,” *Adv. Data Anal. Classif.*, vol. 18, no. 2, pp. 325–361, 2024, doi: 10.1007/s11634-023-00536-8.
[12] S. B. Deshpande *et al.*, “Elevating educational insights: sentiment analysis of faculty feedback using advanced machine learning models,” *Adv. Contin. Discret. Model.*, vol. 2025, no. 1, 2025, doi: 10.1186/s13662-025-03933-9.
[13] N. M. S. Iswari, “Analyzing Student Sentiments and Insights on Generative AI for Independent Learning in Universities,” *J. Appl. Data Sci.*, vol. 7, no. 1, pp. 786–801, 2026, doi: 10.47738/jads.v7i1.1083.
[14] D. Baidoo-Anu and L. O. Ansah, “Education in the Era of Generative Artificial Intelligence (AI): Understanding the Potential Benefits of ChatGPT in Promoting Teaching and Learning,” *DergiPark (Istanbul Univ.*, vol. 7, no. 1, pp. 52–62, 2023, doi: 10.61969/jai.1337500.
[15] D. K. Kanbach, L. Heiduk, G. Blueher, M. Schreiter, and A. Lahmann, “The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective,” *Rev. Manag. Sci.*, vol. 18, no. 4, pp. 1189–1220, 2023, doi: 10.1007/s11846-023-00696-z.
[16] N. Rane, S. Choudhary, and J. Rane, “Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation,” *J. Appl. Artif. Intell.*, vol. 5, no. 1, pp. 69–93, 2024, doi: 10.48185/jaai.v5i1.1052.
[17] H. Pratiwi, M. I. Sa’ad, N. Hikmah, and A. Prima, “Integrating Artificial Intelligence in Formative Assessment: Connecting Student Engagement, Learning Styles, and Learning Outcomes,” *J. Pedagog. Educ. Sci.*, vol. 5, no. 1, pp. 190–206, 2026, doi: 10.56741/IISTR.jpes.001513.
[18] Q. Xia, X. Weng, F. Ouyang, T. Lin, and T. K. F. Chiu, “A scoping review on how generative artificial intelligence transforms assessment in higher education,” *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, 2024, doi: 10.1186/s41239-024-00468-z.
[19] B. G. A. Enríquez, M. A. A. Ballesteros, O. H. Jordan, C. L. Roca, and K. S. Tirado, “Analysis of college students’ attitudes toward the use of ChatGPT in their academic activities: effect of intent to use, verification of information and responsible use,” *BMC Psychol.*, vol. 12, no. 1, p. 255, 2024, doi: 10.1186/s40359-024-01764-z.
[20] N. Punetha and G. Jain, “Advancing sentiment classification through a population game model approach,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–27, 2024, doi: 10.1038/s41598-024-70766-z.
[21] S. Saepudin, S. Widiastuti, and C. Irawan, “Sentiment Analysis of Social Media Platform Reviews Using the Naïve Bayes Classifier Algorithm,” *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 2, pp. 236–243, 2023, doi: 10.32736/sisfokom.v12i2.1650.
[22] K. Walji, A. Erraissi, A. ZAKRANI, and M. Banane, “From Review to Practice: A Comparative Study and Decision-Support Framework for Sentiment Classification Models,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 9, pp. 699–709, 2025, doi: 10.14569/IJACSA.2025.0160967.
[23] M. A. Talukder *et al.*, “A hybrid deep learning model for sentiment analysis of COVID-19 tweets with class balancing,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–19, 2025, doi: 10.1038/s41598-025-97778-7.
[24] R. ElRabaa, S. Rawas, and A. El-Zaart, “A Gamma-enhanced Naïve Bayes model for robust intrusion detection in IoMT networks,” *Discov. Internet Things*, vol. 6, no. 1, 2026, doi: 10.1007/s43926-025-00281-8.
[25] A. Almassaad, H. Alajlan, and R. Alebaikan, “Student Perceptions of Generative Artificial

Intelligence: Investigating Utilization, Benefits, and Challenges in Higher Education," *Systems*, vol. 12, no. 10, p. 385, 2024, doi: 10.3390/systems12100385.