Isabela State University Linker:
Journal of Engineering, Computing, and Technology
Volume 2, Issue 2
ISSN (Print): 3082-3676 ISSN (Online): 3082-3684
DOI: https://doi.org/10.65141/ject.v2i2.n4

Sentiment Analysis and Word Cloud of Teachers' Evaluations Using R Programming Language

Catleen Glo M. Feliciano[1]
College of Computing Studies Information and Communication Technology, Isabela State University, Echague, Isabela, 3309, Philippines[1]
catleenglo.r.madayag@isu.edu.ph

RESEARCH ARTICLE INFORMATION

Received: July 10, 2025
Reviewed: November 13, 2025
Accepted: November 18, 2025
Published: December 31, 2025

ABSTRACT

Faculty evaluation is essential for understanding students' perceptions and feedback to improve the employment of teaching strategies. With the use of vast-scale textual feedback, in an efficient manner, sentiment analysis was used as a tool for analyzing textual semantics in a structured way that could help facilitate understanding of what students think. Using the datasets of students' feedback from faculty evaluation from A.Y. 2019-2020 to A.Y. 2024-2025 for sentiment analysis using R programming, this study utilized Natural Language Processing (NLP). Data preprocessing, word cloud creation, and sentiment classification using code were employed to systematically extract prevalent themes, classify sentiments, and examine faculty performance. The approach comprises several processes, such as data preprocessing, word cloud generation, and sentiment classification, which are used to classify sentiments that follow an organized topic extraction and present useful insights about teacher performance. In fact, according to the data, students are overwhelmingly positive, with a deep appreciation for teachers who are helpful, efficient, and supportive in their teaching style and approach. The result also reflects how much students value the hard work that their teachers do, such as the top positive word is kind (*mabait*). Though they are less common, unfavorable opinions do draw attention to the areas in which students struggle, especially when it comes to their academic performance. While there are terminologies that reflect occasional problems in the classroom, where the top negative words are limit and hardship (*hirap*), it was noted that certain students struggle with their tasks. The results highlight how crucial it is to have a welcoming and interesting learning environment. Teachers may reinforce their strengths and highlight areas for growth by using sentiment analysis to get insightful information about student responses. Finally, by ensuring a well-rounded, efficient, and student-centered teaching approach for students pursuing a Bachelor of Science in Computer Science, this study offers a data-driven method of improving the learning experience.

Keywords: *big data analysis, computer science, faculty evaluation, r programming language, sentiment analysis, word cloud*

Introduction

In state universities and Colleges, faculty member evaluations of performance are important since they provide ideas regarding students' opinions and enhance instructional strategies (Delgado & Cabiles, 2024; Chaudhry et al.,2023; Kim et al., 2024). Due to the continuing improvement of technology, text mining techniques such as sentiment analysis and word cloud visualization made it easier to analyze vast amounts of student comments to faculty (Sweta,2024; Takaki & Dutra,2023).

Through faculty evaluation, the students' opinions, attitudes, and views about the faculty were collected (Payandeh et.al., 2023). The evaluations include useful data that is used to evaluate the effectiveness of instruction and identify areas in need of improvement (Constantinou & Wijnen-Meijer, 2022).

Faculty evaluations do more than measure teaching performance (Facciolo & Pittenger, 2024). It also helps schools improve the learning environment (Arifin et.al., 2024). Sentiment analysis supports this goal by giving a clear view of what students feel and think (Grimalt & Usart, 2024). It can reveal common concerns, show patterns over time, and highlight areas that need attention (Salgado et.al.,2024; Sharma et.al., 2025). These insights can guide faculty development programs, improve classroom practices, and support curriculum updates (Deshpande et.al., 2025). With these benefits, sentiment analysis becomes a useful tool for improving both teaching quality and academic decision-making.

On the other hand, Natural Language Processing (NLP) is a field of artificial intelligence that allows computers to understand and analyze human language by converting unstructured text into a form that can be processed and interpreted by a machine (Akhil et al., 2024).

In this study, NLP techniques were used to clean the student comments, remove noise, standardize words, and prepare them for analysis, making it possible to detect patterns, identify sentiment, and extract keywords from a large set of faculty evaluation comments written in English and Tagalog. As textual feedback is unstructured, it might be challenging to derive students' perception from it (Lin et al., 2025), and thousands of written comments may be included in a single dataset, making it difficult to manually evaluate and understand the sentiment as a whole (Jim et al., 2024).

This study analyzed and represented student input by utilizing word clouds and sentiment analysis tools, exposing trends in students' opinions. The main objectives of this study were to apply techniques from natural language processing in sentiment analysis, clean and structure unstructured faculty evaluation data, generate word clouds to help visualize significant subjects in student comments, categorize feelings into negative or positive groups, and provide views that enable teachers to improve their strategies of instruction.

This study used faculty evaluation data from Bachelor of Science in Computer Science students to discuss the text preparation, sentiment analysis, and word cloud creation processes. The objective was to discover commonly used terms that reflect students' experiences and determine if the majority of their feedback is positive or negative.

While some students used English, most of the comments were written in Tagalog. The Tagalog remained the most common language in the comments as it was the main language spoken by the students (Balahadia et.al., 2016) at Isabela State University, thus the researcher considered it in doing the sentiment analysis. There were several challenges faced during the conduct of the study, such as unstructured data where student comments are frequently informal and may contain wording variations, making standardization challenging; and context interpretation, where certain words may have different meanings depending on the context, requiring careful analysis. To help improve teaching methods and the learning environment, this study was conducted to address these issues and advance a more thorough knowledge of faculty performance based on student comments.

Many studies on sentiment analysis focus on English texts and structured survey responses (Yacoub et.al., 2024), but few examined comments written in more than one language (Sharma et.al., 2025), such as the mixed Tagalog–English feedback common in Philippine SUCs. This lack of research creates problems in text cleaning, sentiment tagging, and word interpretation because existing tools often rely on English-only datasets, simple lexicons, or pre-trained models that do not match the language used by students.

As a result, current methods may fail to capture the true meaning and tone of student comments. This study addresses the gap by using NLP techniques that work with both Tagalog and English, account for spelling variations, and adapt lexicon-based and sentiment methods to a multilingual setting. Through this approach, the study offers a more accurate and relevant way to analyze faculty evaluation feedback in environments where code-mixed language is natural and widely used.

Table 1 presents a comparison table that shows that the study addresses the gap by customizing preprocessing, lexicons, and sentiment techniques specifically for code-mixed Tagalog–English faculty evaluation comments, something not well-covered in existing literature.

Table 1. Tabular Comparison of Methods for Multilingual Sentiment Analysis

Method	Description	Data Requirement	Expected Accuracy	Strength
Lexicon-Based Methods	Uses predefined lists of positive and negative words	Very low; can work with small datasets	Low–Moderate	Simple, fast, interpretable; no training needed
Traditional Machine Learning	Classifies text using features such as bag-of-words or TF-IDF	Moderate; requires labelled training data	Moderate	Works well with small to medium datasets; adaptable
Deep Learning	Learns patterns automatically from large datasets	High; needs many labelled examples	Moderate–High	Captures complex linguistic patterns; better context understanding
Multilingual Transformer Models	Uses pre-trained multilingual language models for sentiment analysis	Medium–High; benefits from fine-tuning	High	Best performance with mixed languages; handles context, grammar, and slang effectively

Methods

Improving teaching methods in the 21st century requires an awareness of how students view teacher performance (Poonputta & Nuangchalerm, 2024). In order to better understand students' opinions, this study examined student comments from instructors' evaluations to determine students' opinions, feelings, and assessments based on their written input by using sentiment analysis and word cloud visualization. Natural Language Processing (NLP) is a branch of artificial intelligence that enables computers to process, analyze, and interpret human language (Rongali, 2025).

In this study, NLP techniques were employed to convert unstructured student comments from faculty evaluations into structured data that could be quantitatively analyzed. NLP preprocessing allowed for the identification of patterns, extraction of meaningful words, and sentiment classification (Dogra et.al., 2022) in both English and Tagalog comments.

Research Design

This study analyzed student feedback from instructor evaluations using a quantitative text mining technique that makes use of word cloud visualization and sentiment analysis (Chavan et.al., 2024). While word cloud visualization reveals often used terms (Skeppstedt et.al., 2024) and offers insights into how students view their teachers, sentiment analysis categorizes the input into positive and negative attitudes (Hasan et.al., 2024). Data gathering, preprocessing, sentiment analysis, and R programming data visualization are all steps in the study's organized research methodology (Shahare et.al.,2024).

The quantitative text mining process follows a clear and structured workflow. First, raw student comments are collected and imported into the system. Next, NLP preprocessing steps—including text cleaning, tokenization, stop-word removal, and word normalization—are applied to prepare the data for analysis. The cleaned text is then processed using two main techniques: sentiment analysis, which determines whether each comment conveys a positive or negative tone, and word cloud generation, which highlights the most frequently used terms.

Finally, the results are summarized and visualized through word clouds, bar plots, and sentiment distribution charts to provide meaningful insights into faculty performance. This architecture ensures a systematic transition from unstructured text to interpretable and actionable findings.

Participants and Locale of the Study

Participants in this study included Bachelor of Science in Computer Science students from Isabela State University-Echague Campus. Students' comments on instructor evaluations from the first semester of A.Y. 2019–2020 were the dataset for the first semester of A.Y. 2024–2025. Direct contact with respondents was not necessary because the emphasis of this study is on textual data analysis. Rather, analysis was done using the faculty evaluation data that was already available.

Research Instruments

Comments from instructors were used as the primary source of data for this study. The sentiment analysis and text mining process were carried out using R programming, especially making use of tools like:

tidytext: for tokenization and text preprocessing
Wordcloud: for a graphic representation of word frequency
Syuzhet: for sentiment analysis
ggplot2: for data visualization

The study utilized R programming with libraries including tidytext, tm, syuzhet, wordcloud, and ggplot2 due to their strengths in text mining, sentiment analysis, and data visualization. The tidytext enabled tokenization and efficient manipulation of textual data in a tidy data framework, making it straightforward to integrate with dplyr for preprocessing. Syuzhet provided lexicon-based sentiment scoring and easily accommodated custom lexicons, which is crucial for analyzing both English and Tagalog comments. The wordcloud and wordcloud2 libraries allowed flexible and visually informative representation of frequently used words. Compared with alternative tools such as Python’s NLTK or TextBlob, R was chosen for its seamless integration with data visualization, reproducibility in academic reporting, and the availability of libraries suitable for multilingual lexicon-based sentiment analysis.

Data Collection Procedures

This section describes the methodical procedure of the study for collecting, classifying, and evaluating student comments to do sentiment analysis utilizing both Tagalog and English lexicons. The data used were extracted from the university's faculty evaluations spanning five academic years. The study gathered student comments and compiled them into one .csv format file, with a total of 605 student comments collected. The data were cleaned by removing special characters, stop words, and redundant keywords. Using an .xlsx file containing a custom Tagalog sentiment lexicon that gives sentiment scores to words and the AFINN lexicon for English terms, wherein the cleaned dataset was used for a custom lexicon-based analysis.

The dataset consisted of student comments extracted from the university's faculty evaluation system, covering the first semester of A.Y. 2019–2020 to the first semester of A.Y. 2024–2025. All personally identifiable information was removed, and the data were de-identified before analysis. Since the study involved retrospective analysis of institutional data, direct consent from students was not required; however, the project followed the university's research ethics guidelines. The dataset is retained securely in CSV and Excel formats, accessible only to the college program chairs, and will be maintained for reproducibility for five years.

Data Cleaning Details

To ensure reproducibility, the methodology now includes specific steps for data cleaning. The dataset was first consolidated into a single .csv file and inspected for duplicates and empty entries, which were removed to prevent redundancy. Special characters such as punctuation marks (!, .,,, ?, etc.) and numerals were stripped using the tm and stringr libraries. Words were converted to lowercase to standardize the text, and extra whitespace was removed. Stopwords in English and Tagalog—common but semantically weak words such as "the," "and," "si," and "ng"—were removed to reduce noise.

Additionally, redundant or repeated keywords were filtered to avoid overrepresentation in word frequency analysis. These steps ensured that only meaningful terms were retained for tokenization, sentiment analysis, and word cloud generation. For reproducibility, the R code snippets using tm_map, removePunctuation(), removeNumbers(), and custom stopword lists were added in the revised methodology.

Text Preparation

The text preparation stage involved removing unnecessary characters, correcting spelling variations, converting text to lowercase, and separating each word. Stop words in both English and Tagalog were removed to keep only meaningful terms. After cleaning, sentiment analysis was performed by matching each word with a lexicon containing positive and negative Tagalog and English terms.

The system then assigned a sentiment score to each comment based on the words it contained. For the word cloud, the cleaned words were counted and arranged based on their frequency, allowing common themes in the comments to become visible. These processes helped reveal the tone, patterns, and key topics in student feedback.

Analysis of Data

The qualitative interpretation of the textual data and descriptive statistics was used in the study. Sentiment scores were computed by classifying words into three categories: neutral, negative, and positive (Raees & Fazilat, 2024). Finding commonly used terms that reflected students' overall perceptions of their teachers was made easier by the word cloud display. Additionally, bar graphs were created to show the distribution of sentiment throughout the dataset.

The methodology addresses potential bias by including all student comments over five academic years, without manual selection, to reduce sampling bias. Comments were first identified as English or Tagalog based on the predominant language of the words, using the merged lexicon for token comparison. Sentiment scoring was performed at the word level: each token was matched against the AFINN lexicon for English and the custom Tagalog lexicon. Words were assigned sentiment values (+1 for positive, -1 for negative, 0 for neutral), and the overall comment score was calculated as the sum of its constituent words' scores.

This approach ensures consistent and transparent classification of comments into positive, negative, or neutral categories. To further reduce bias from code-mixed comments, combined lexicons, and stop words in both languages were applied during preprocessing and scoring.

To validate the sentiment analysis, approximately 10–20% of the comments were randomly sampled and manually labeled as positive, negative, or neutral. Inter-annotator agreement was computed using Cohen's κ to ensure consistency between multiple human reviewers. The performance of the lexicon-based sentiment scoring was quantified using precision, recall, and F1-score metrics. Errors were analyzed to identify common misclassifications, particularly in code-mixed Tagalog–English comments.

To improve sentiment detection, the methodology was enhanced with bigrams and trigrams, e.g., no feedback, so hard (*sobrang hirap*); negation scope handling, e.g., unkind (*hindi mabait*); intensifiers, e.g., so (*sobrang*), very; and polarity shifters to accurately reflect meanings in both English and Tagalog. These enhancements help capture context-dependent sentiment that single-word analysis may miss. The custom Tagalog–English lexicon was benchmarked against AFINN on the English subset to evaluate coverage and accuracy.

For future work, the study proposes the integration of refined multilingual transformer models, such as mBERT or XLM-R, to further improve sentiment classification on code-mixed comments and address limitations of lexicon-based methods.

The sentiment analysis used in this study classified the comments into two main classes: positive and negative. These classes were based on a lexicon of English and Tagalog words with assigned sentiment values. Positive words included terms related to helpfulness, clarity, and effective teaching, while negative words included terms linked to poor communication, unclear instruction, or undesirable behavior. The classification made it possible to measure the overall emotional tone of the feedback and determine whether the majority of student comments expressed approval or concern.

Research Process

This section outlines the procedures that the researcher uses to generate sentiment analysis and word clouds. The R Studio IDE and the R programming language were utilized by the researcher.

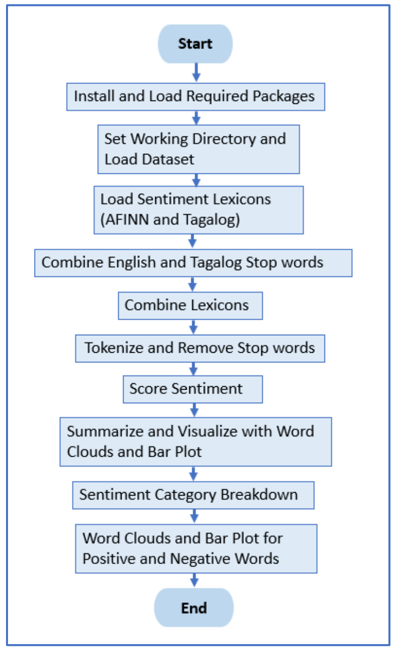

Figure 1. *The Steps of Analysis*

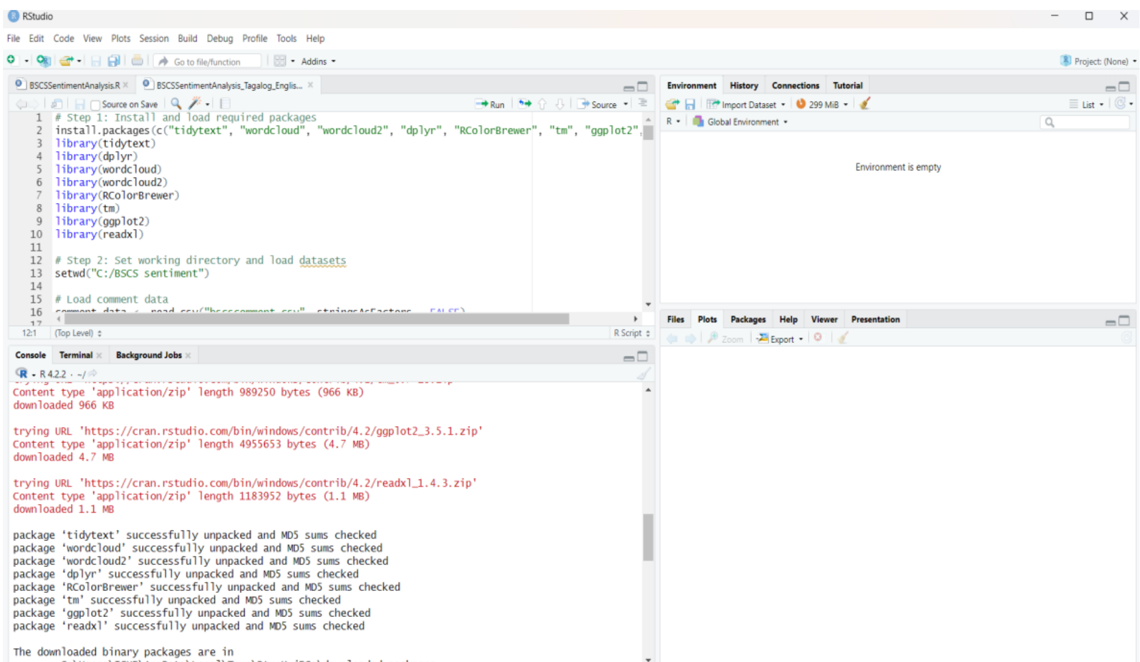

Figure 2. *Install and Load Required Packages*

The programming environment was prepared by installing and loading the necessary packages for text mining (tidytext, tm), data manipulation (dplyr), visualization (ggplot2, wordcloud, wordcloud2), reading files (readxl), and applying color themes (RColorBrewer).

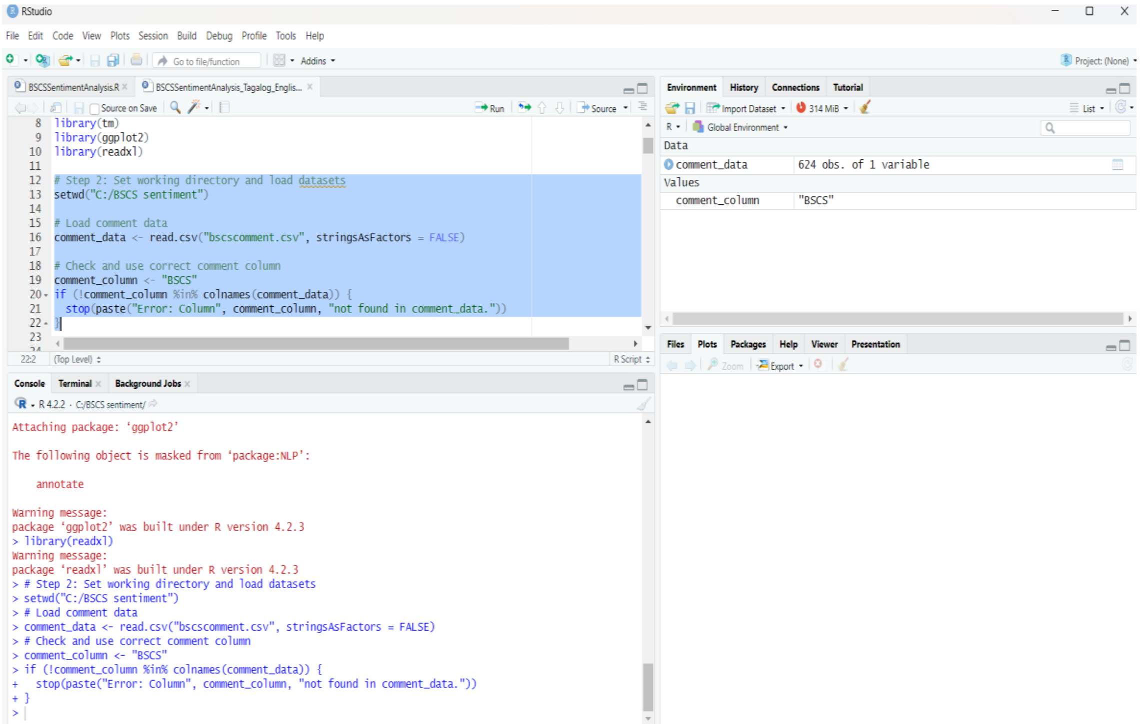

Figure 3. *Set Working Directory and Load Dataset*

The method for loading the dataset with student comments for processing is illustrated in Figure 3.

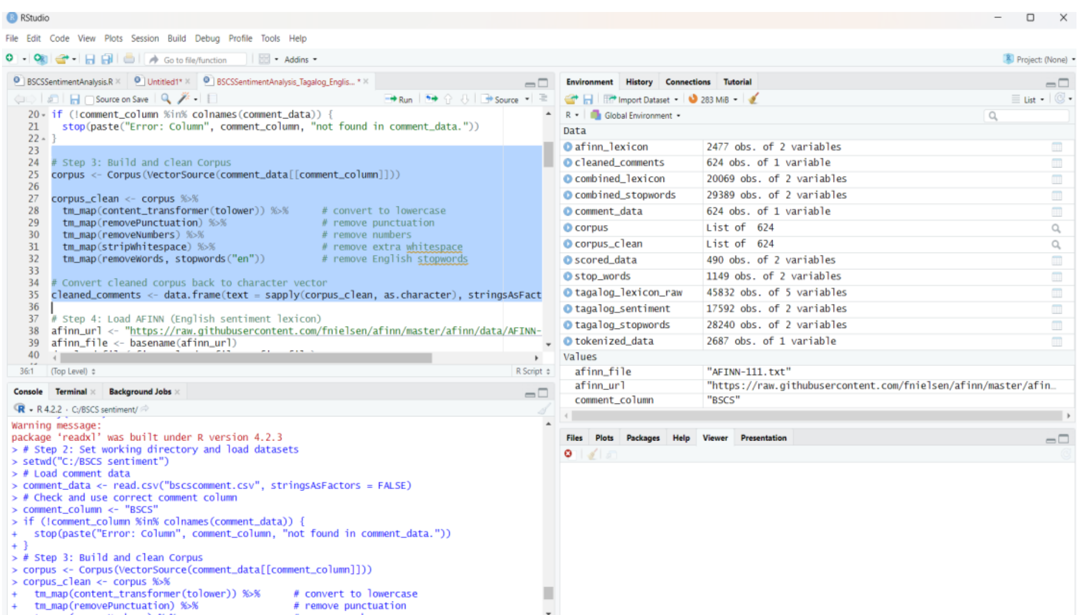

Figure 4. *Build a Corpus*

The construction of a text corpus for Natural Language Processing (NLP) tasks is shown in Figure 4. The feedback comments were converted into a structured format for this corpus, which made manipulation and analysis simpler. After that, all words were changed to lowercase, punctuation was removed, numerals were removed, extra whitespace was removed, and frequent stop words in both English and Tagalog were filtered out of the text data.

To reduce noise and make sure that only relevant words contribute to the analysis, stop words—words that do not convey substantial meaning—such as "the," "and," "si," and "ng"—were eliminated. To prepare it for other NLP processes like tokenization, sentiment analysis, and word frequency analysis, the cleaned corpus was then transformed into a character vector.

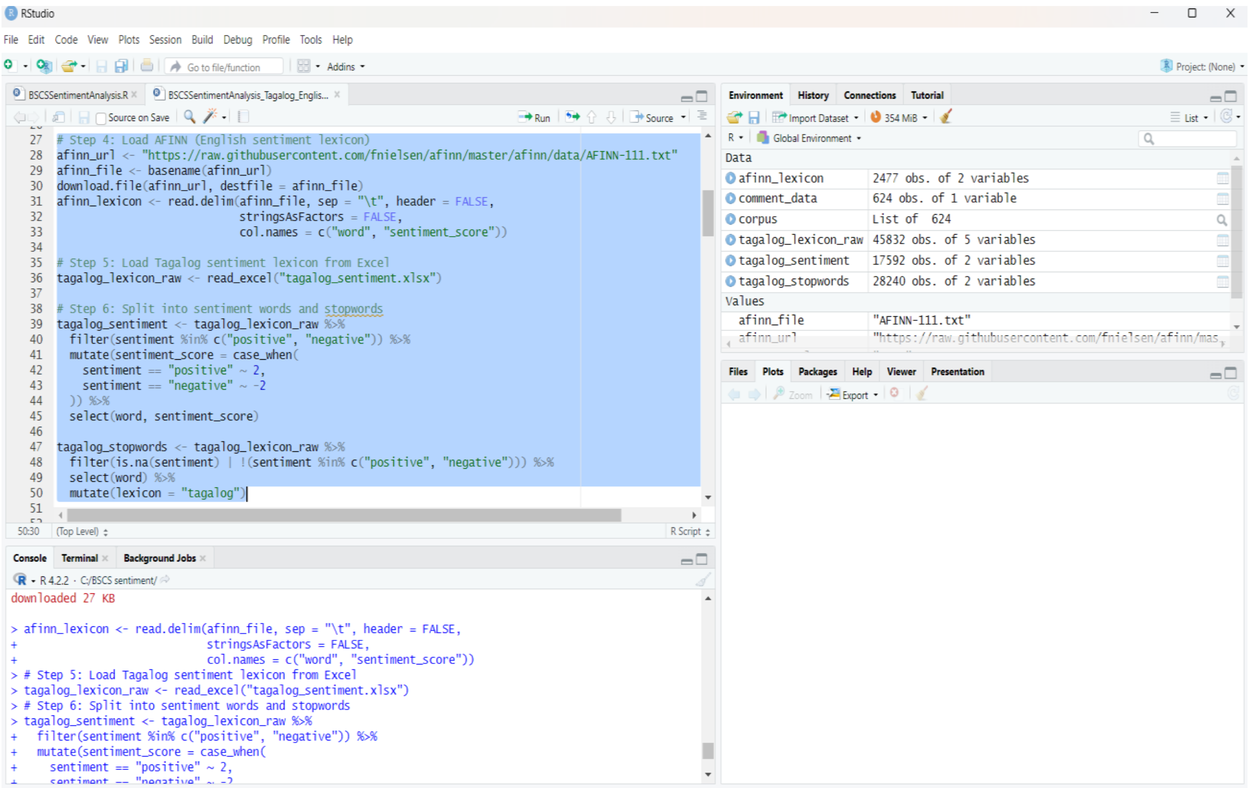

Figure 5. *Load Sentiment Lexicons (AFINN and Tagalog)*

The use of two emotion lexicons is illustrated in Figure 5, where AFINN for sentiment scoring in English and a customized Excel lexicon included both positive and negative evaluations for Tagalog words.

On the other hand, Figure 6 illustrates how common or unnecessary terms from both languages—such as "is", "in", "to", "the", "at", "ang", "at", "sa", etc.—were eliminated throughout the analysis process.

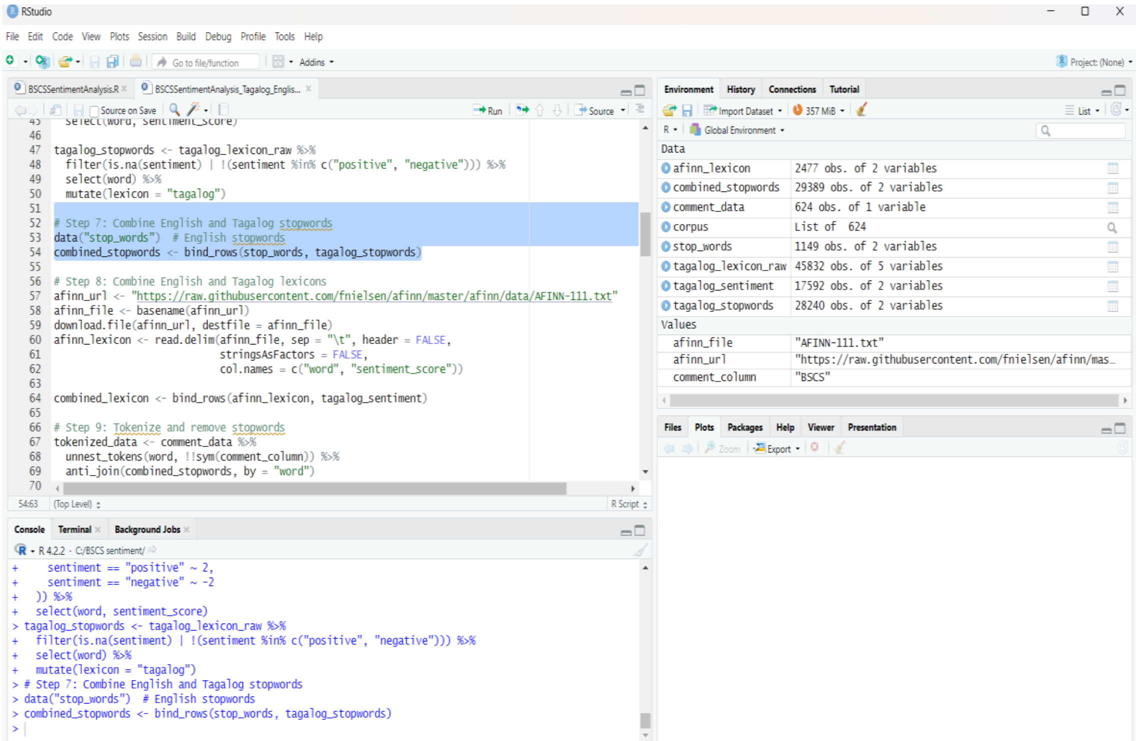

Figure 6. *Combining English and Tagalog Stop Words*

In addition, Figure 7 establishes a common vocabulary for sentiment analysis in Tagalog and English.

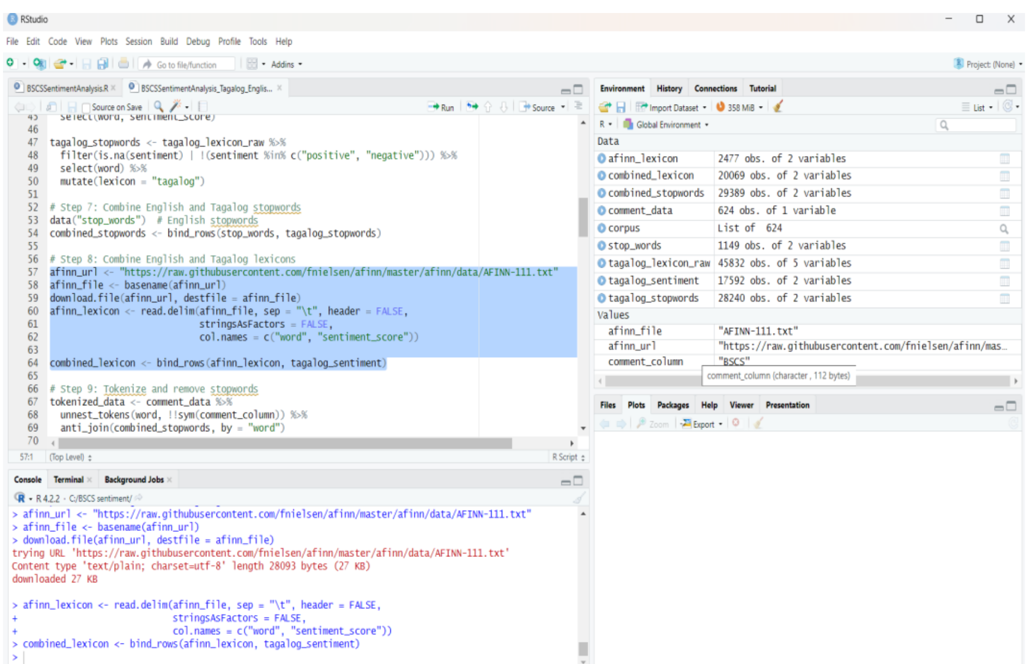

Figure 7. *Lexicon Combination*

To identify important words, break phrases up into their individual words (tokens) and eliminate stop words (See Figure 8).

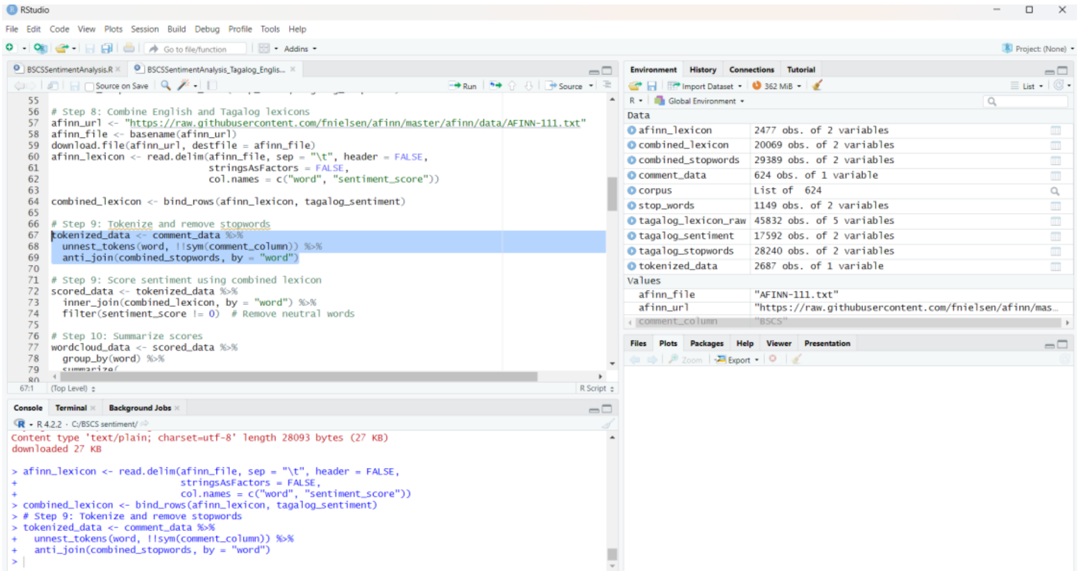

Figure 8. *Tokenization and Removal of Stop Words*

Based on the merged lexicon, give words sentiment ratings (Refer to Figure 9).

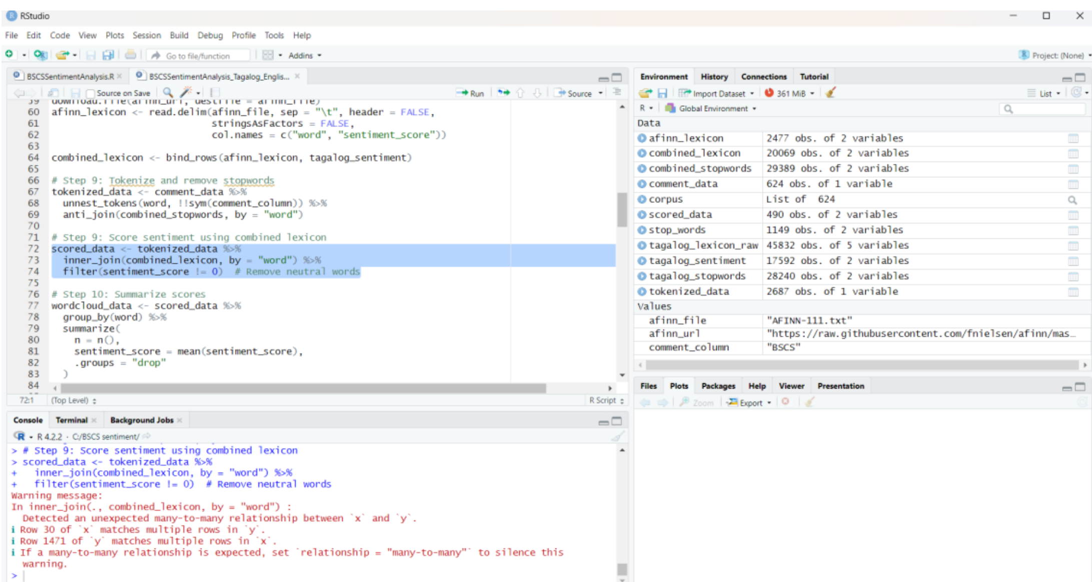

Figure 9. *Score Sentiment*

Figures 10, 11, and 12 show how the wordcloud2 function is used to visualize common sentiment-bearing words in various forms (triangle, star, and circle), as demonstrated by the code wordcloud2 (wordcloud_data, shape = "circle"), wordcloud2(wordcloud_data, shape = "star"), and wordcloud2(wordcloud_data, shape = "triangle")

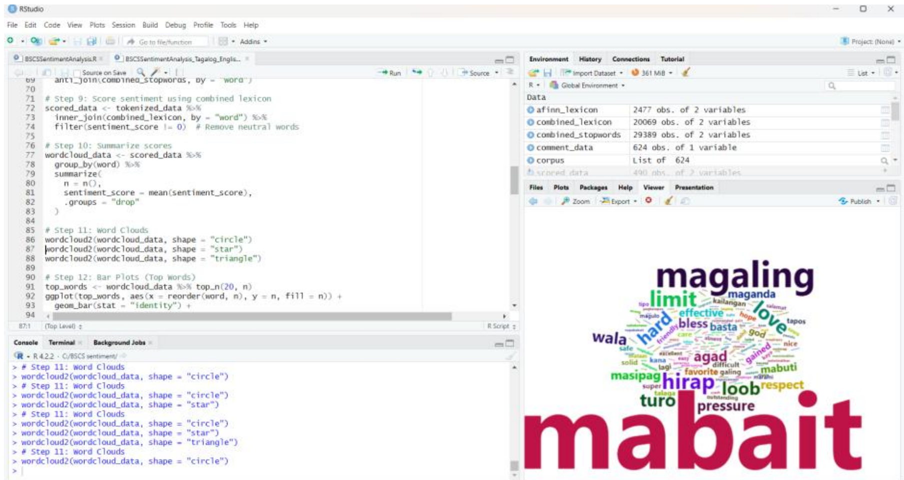

Figure 10. *Summarization and Visualization Through Word Clouds Circle Shape*

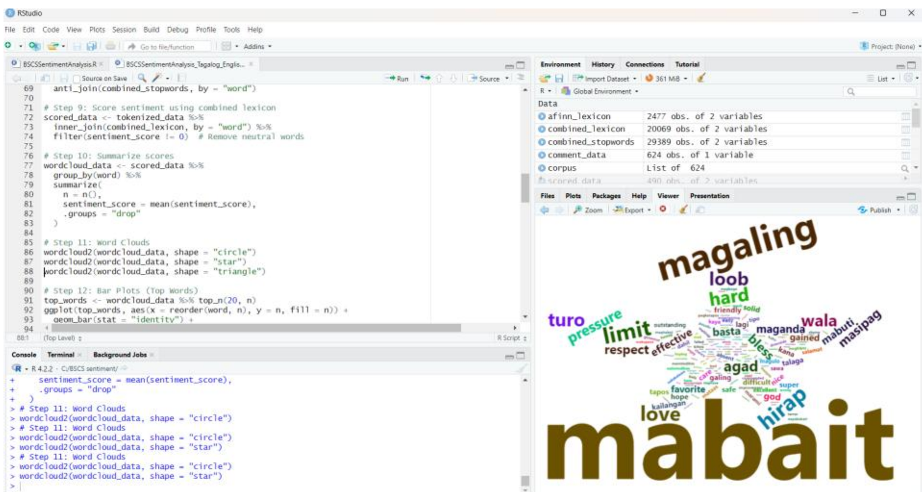

Figure 11. *Summarization and Visualization Through Word Clouds Star Shape*

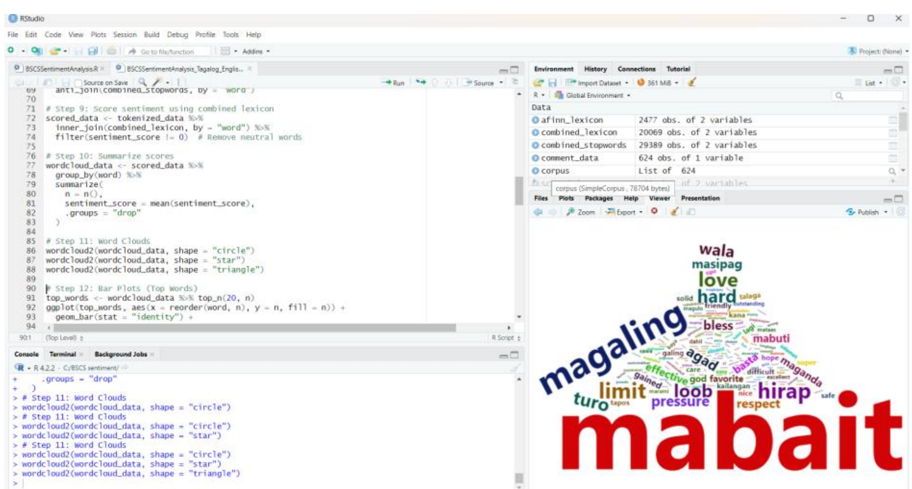

Figure 12. *Summarization and Visualization with Word Clouds Triangle Shape*

Figure 13 displays the creation of a bar chart that shows the most common terms from student comments.

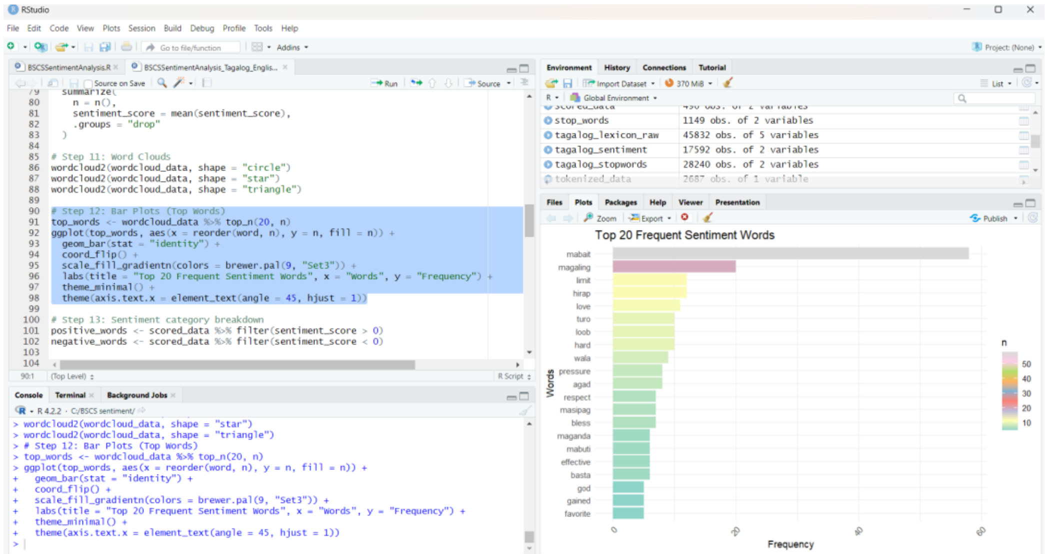

Figure 13. *Bar Plot of Top Words*

Figure 14 shows a method for separating positive and negative feedback and determining their respective proportions.

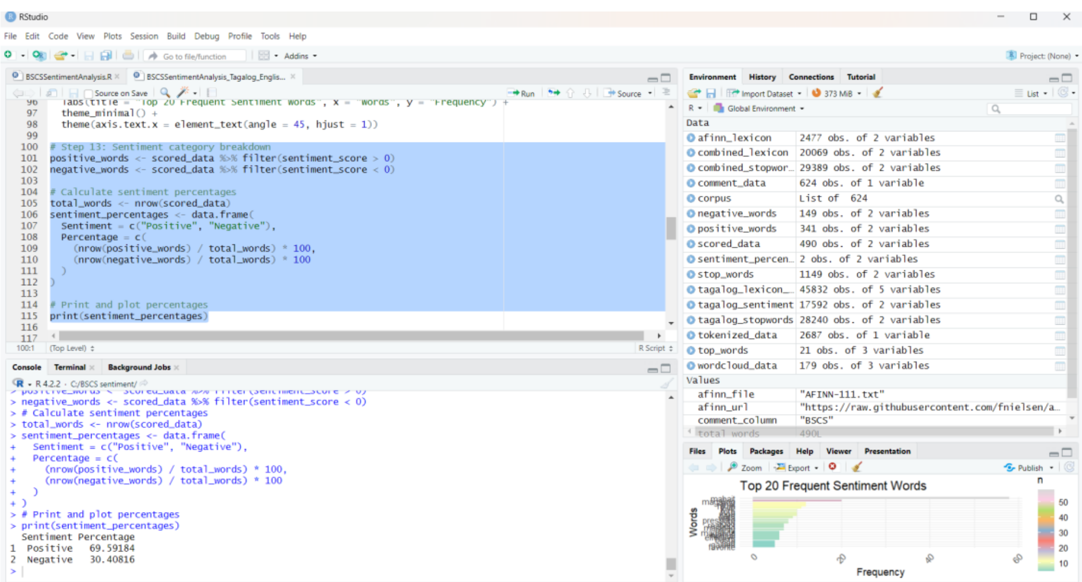

Figure 14. *Sentiment Category Breakdown*

The following figures (Figures 15 to 18) demonstrate how to make unique word clouds and bar plots for positive and negative terms to visualize sentiment distinctly.

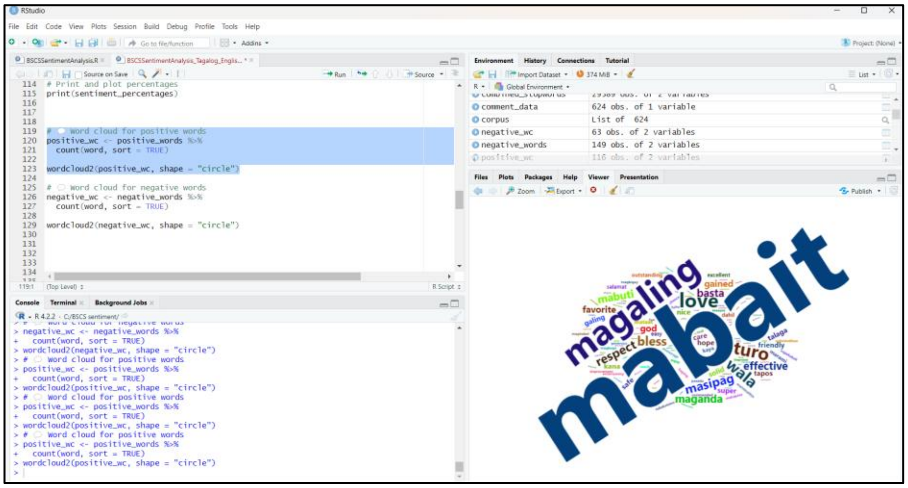

Figure 15. *Word Clouds for Positive Words*

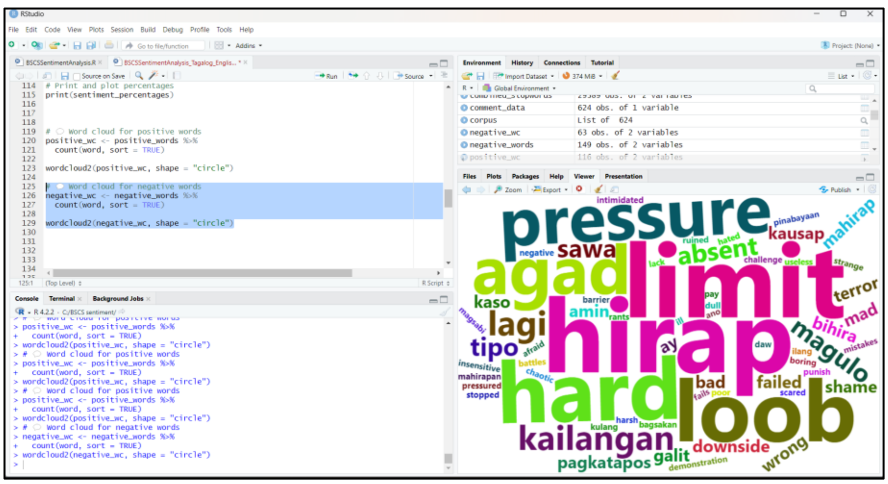

Figure 16. *Word Clouds for Negative Words*

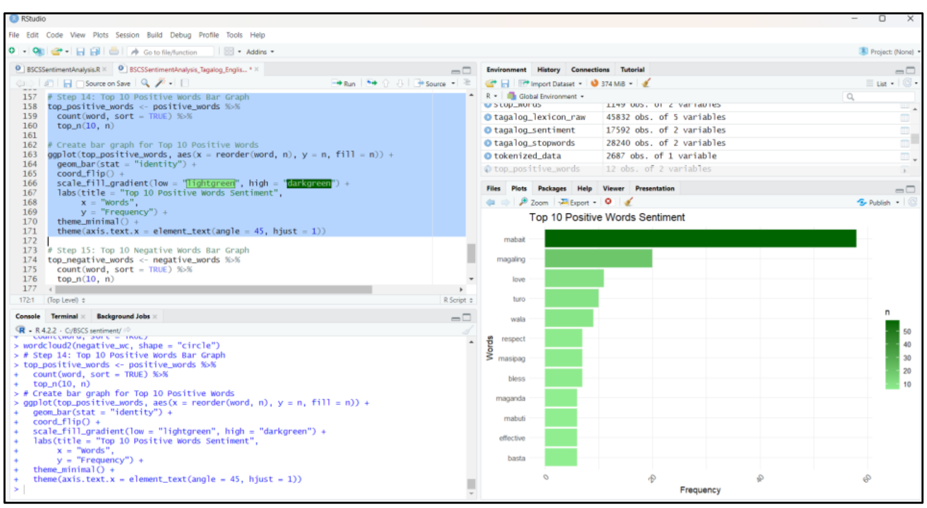

Figure 17. *Bar Plot for Top 10 Positive words*

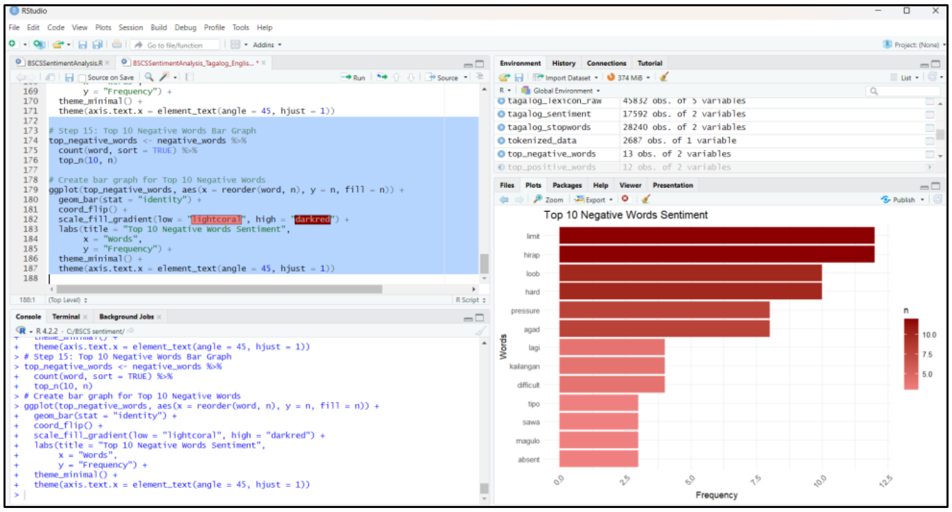

Figure 18. *Bar Plot for Top 10 Negative Words*

Ethical Considerations

This study ensured the confidentiality and anonymity of student responses by closely adhering to ethical research requirements. No personally identifiable information was utilized, and an aggregate analysis of the dataset was conducted. The study was carried out in accordance with institutional guidelines for ethical research procedures and data protection.

Results and Discussion

The study focused on the application of NLP techniques to student comments from faculty evaluations through the R programming language and R Studio IDE to analyze comments. The findings provide an understanding of students' perspectives, highlighting both positive and negative sentiments toward the faculty. To guarantee specific sentiment categorization, the input data was preprocessed by removing special characters, stop words, and unnecessary keywords. After cleaning, the data was organized for sentiment analysis and word cloud creation. The analytic stages are listed below.

Table 2. Code Mapping

Research Objective	Code Steps Addressing It
Apply NLP techniques	Steps 1, 3, 4–9
Clean and structure data	Steps 2, 5–9
Generate word clouds	Steps 11, 14
Classify sentiments	Steps 4–10, 13
Provide insights for teaching	Steps 12, 13, 14

Table 2 summarizes how the study's research objectives were addressed through specific steps in the R-based methodology. For instance, the application of NLP techniques was implemented in Steps 1, 3, and 4–9, which included loading packages, creating a text corpus, cleaning data, tokenizing comments, and preparing sentiment lexicons. The objective of cleaning and structuring unstructured faculty evaluation data was accomplished in Steps 2 and 5–9, covering data loading, preprocessing, stop word removal, and combining English and Tagalog lexicons. Generating word clouds (Steps 11 and 14) provided a visual representation of frequent sentiment-bearing words, while sentiment classification (Steps 4–10, 13) allowed the identification of positive and negative comments. Finally, Steps 12, 13, and 14 enabled the study to provide actionable insights for teaching by summarizing sentiment distributions and highlighting key areas for instructional improvement. This mapping demonstrated the systematic alignment of coding steps with research objectives, ensuring that each goal—from data preparation to actionable insights—was methodically achieved.

Application of Natural Language Processing (NLP) Techniques

To ascertain the sentiment included in a collection of comments from BSCS students, natural language processing (NLP) techniques were applied. Because these comments were written in both Tagalog and English, it was necessary to incorporate multilingual techniques and resources into the analysis to guarantee a thorough assessment of opinions.

To facilitate word-level analysis, the approach started with text preprocessing, in which student comments were tokenized—that is, divided into individual words. Both Tagalog stop words from a bespoke lexicon and English stop words from the tidytext package were used in the combined stop word list. Words like "ang," "ng", "the", and "is" which do not add much to sentiment analysis, were eliminated in this phase.

Two sentiment lexicons were used to take into consideration the multilingual character of the responses: an organized Tagalog sentiment lexicon, in which words were manually classified as either positive or negative, and the AFINN lexicon for English sentiment scoring. An emotion score—+2 for positive and -2 for negative—was given to each word in the Tagalog language. The tokenized words were then scored using a combined sentiment dictionary that was created by combining these two lexicons.

To make sure that only content that expressed an opinion was examined, words that were not included in either lexicon were not included in the sentiment assessment. Sentiment ratings reflecting the emotional tone of the student input were assigned by matching each remaining word with the combined lexicon.

Word clouds that highlighted the frequency and emotional context of both positive and negative words were created independently in order to display the results. Bar charts offered a clear, proportionate representation of the distribution of positive and negative attitudes across the dataset and also showed the sentiment-bearing phrases that were used the most.

The common opinions of BSCS students were successfully revealed by this NLP-driven method. Commonly used positive terms that conveyed gratitude and satisfaction, as well as negative words that suggested areas that would want development, were found through the study. Through the integration of bilingual sentiment lexicons, language-aware preprocessing, and straightforward data visualizations, the technique provided a strong and perceptive way to examine student comments. Making well-informed decisions on curriculum development, instructional strategies, and program improvement in general can be aided by these insights.

Given the bilingual nature of the dataset, context interpretation posed challenges (Balahadia & Commendador, 2016). Some words carry different meanings depending on the language; for example, "*loob*" in Tagalog can imply an internal state or attitude, while English words like "*love*" directly express positive sentiment. To address this, separate lexicons were used for each language, and ambiguous or context-dependent terms were manually reviewed. Language-aware tokenization, stop word removal for both languages, and normalization of spelling variations further ensured accurate sentiment classification. These measures minimized misclassification and allowed the analysis to reflect students' intended meanings accurately.

Clean and Structure Unstructured Faculty Evaluation Data

To prepare the unstructured faculty evaluation comments for analysis, the raw dataset was first loaded and inspected (Step 2). Text preprocessing involved converting all text to lowercase, removing punctuation, numbers, and extra whitespace, and standardizing both English and Tagalog stop words (Steps 5–7). English and Tagalog sentiment lexicons were combined to ensure consistent sentiment scoring across languages (Step 8). Finally, the cleaned comments were tokenized, and non-informative words were removed, resulting in structured word-level data suitable for sentiment analysis (Step 9). This process transformed raw, unstructured feedback into a clean, organized format, enabling reliable NLP-based insights and visualizations in subsequent steps.

Generating Word Clouds for Visualizing Key Themes in Student Feedback

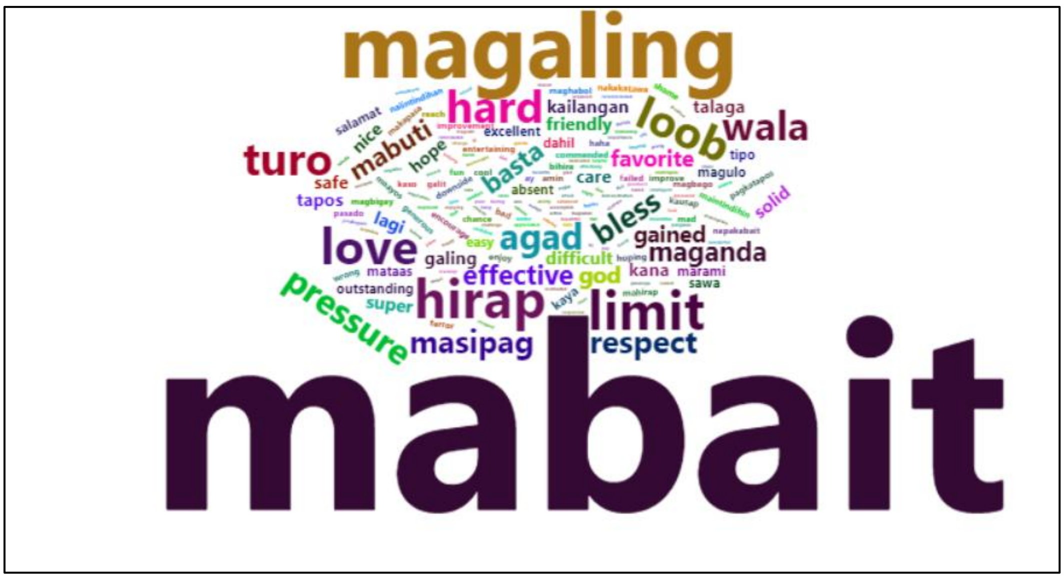

Figure 19. *Word Cloud of Word Sentiment*

Key themes in student response were shown using word clouds, which often highlighted emotional terms in various forms (triangle, star, and circle). With terms like kind (*mabait*), intelligent (*magaling*), love, industrious (*masipag*), good (*mabuti*), and effective, the circle-shaped word cloud highlighted positive comments and expressed gratitude for the instructor's attributes and efficacy as a teacher.

Negative adjectives like hard (*hirap*), none (*wala*), pressure, limit, and teach (*turo*) described difficulties that students experienced. The triangle and star clouds provided more information about the difficulties of the course and the efficacy of instruction. Future enhancements to the course and teaching methodology will be guided by the captivating way in which these visualizations summarize student sentiments.

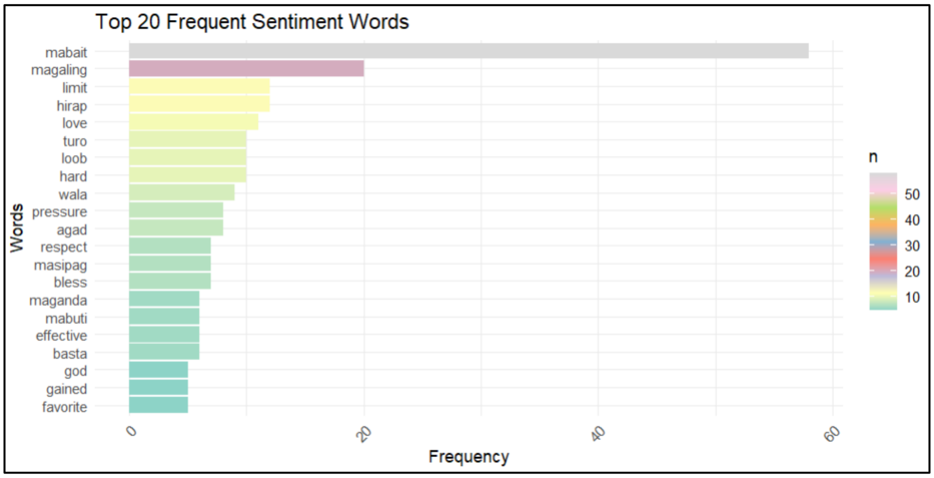

Figure 20. *Bar Plot Top 20 Frequent Sentiment Words*

The most common opinions in student comments are clearly and concisely represented by the bar plot of the top 20 frequently used terms. It shows that 15 of the top 20 words are positive, highlighting how much pupils value their teacher. The list is dominated by positive phrases that reflect the educational approach's characteristics, such as *mabait* (kind), *magaling* (great), love, respect, effective, and *masipag* (industrious). Conversely, derogatory terms like limit, *hirap* (difficult), hard, pressure, and *wala* (none) highlighted areas in which students had difficulties. As the most prominent term, *mabait* (kind) strengthens the generally good feeling. The main topics of the comments were summarized in this bar plot, which offers insightful information for improving the course.

Table 3 presents the top tokens identified in the dataset, including their frequencies, part-of-speech (POS) tags, and sentiment labels. The results show that most of the highly frequent words carry positive sentiment. Tokens such as *mabait* (kind), *magaling* (intelligent), *maganda* (beautiful), and *mabuti* (good) appear often and are tagged mainly as adjectives, suggesting that students used descriptive words to highlight the instructor's positive behavior and teaching performance.

On the other side, action-related words like *turo* (teach), gained, and bless also appear among the top tokens, showing that students recognized helpful teaching practices and personal learning outcomes. In contrast, negative words such as hirap (difficulty), limit, hard, and pressure occur less frequently but point to common challenges experienced by learners, including task difficulty and academic stress. The mixture of positive and negative tokens reflects the range of student experiences. The strong presence of positive descriptors supports the overall positive sentiment trend observed in the study, while the negative terms help identify areas where instructional strategies and course delivery can be improved.

Table 3. Top Tokens with Frequency, POS Tags, and Sentiment

Token	Frequency	POS Tag	Sentiment Score	Sentiment Label
mabait	58	ADJ	2	Positive
magaling	20	ADJ	2	Positive
love	11	NOUN	3	Positive
turo	10	VERB	2	Positive
wala	9	ADV	2	Positive
bless	7	VERB	2	Positive
masipag	7	ADJ	2	Positive
respect	7	NOUN	2	Positive
basta	6	ADV	2	Positive
effective	6	ADJ	2	Positive
mabuti	6	ADJ	2	Positive
maganda	6	ADJ	2	Positive
favorite	5	ADJ	2	Positive
gained	5	VERB	2	Positive
hirap	12	NOUN	-2	Negative
limit	12	NOUN	-2	Negative
hard	10	ADJ	-1	Negative
loob	10	NOUN	-2	Negative
agad	8	ADV	-2	Negative
pressure	8	NOUN	-1	Negative
difficult	4	ADJ	-1	Negative
kailangan	4	VERB	-2	Negative
lagi	4	ADV	-2	Negative
absent	3	ADJ	-2	Negative

Table 4 presents the sentiment analysis of student comments for the first and second semesters. In the first semester, there were 163 positive words and 54 negative words, making up 75.12% and 24.88% of the total 217 words, respectively. In the second semester, positive words decreased to 101, while negative words slightly increased to 46 out of 147 total words, representing 68.71% positive and 31.29% negative.

Table 4. Sentiment Analysis by Semester

Semester	Positive Words	Negative Words	Total Words	Positive Percentage	Negative Percentage
1st Semester	163	54	217	75.12	24.88
2nd Semester	101	46	147	68.71	31.29

These results show that students' feedback was mostly positive in both semesters. However, the positive sentiment was higher in the first semester, and the negative sentiment increased in the second semester. This may indicate emerging challenges or concerns from students as the academic year progresses. Overall, the table highlights changes in student sentiment across semesters and provides insight into areas that may need improvement or support.

Table 5. Sentiment Analysis by Year

Year	Positive Words	Negative Words	Total Words	Positive Percentage	Negative Percentage
1st Year	87	59	146	59.59	40.41
2nd Year	75	23	98	76.53	23.47
3rd Year	73	13	86	84.88	15.12
4th Year	29	5	34	85.29	14.71

Table 5 shows the sentiment of BSCS students' comments across four academic years. Positive words increased steadily from 59.59% in the 1st year to 85.29% in the 4th year. At the same time, negative words decreased from 40.41% to 14.71%. This trend suggests that students' feedback became more positive as they advanced through the program. The decrease in negative sentiment may indicate growing confidence, familiarity with the curriculum, or improved learning experiences. The results reveal a clear upward trend in positive sentiment and a decline in negative sentiment across the four years.

Table 6. Positive Sentiment Proportion by Year Level with Bootstrapped 95% Confidence Interval

Year level	Positive Words	Negative Words	Total Words	Positive Proportion	95% CI (Lower–Upper)
1st Year	87	59	146	0.596	0.514 – 0.678
2nd Year	75	23	98	0.765	0.684 – 0.847
3rd Year	73	13	86	0.849	0.767 – 0.919
4th Year	29	5	34	0.853	0.735 – 0.971

Table 6 presents the sentiment analysis of student feedback across year levels. The proportion of positive words increased from 0.596 in the 1st Year to 0.853 in the 4th Year. Bootstrapped 95% confidence intervals indicate that these estimates are precise, with minimal overlap between 1st and later year levels. Sentiment analysis of student feedback revealed an increasing trend in positive words from BSCS 1st Year (0.596, 95% CI: 0.514–0.678) to 4th Year (0.853, 95% CI: 0.735–0.971). Bootstrapped confidence intervals confirmed the reliability of these estimates. A chi-square test for proportions indicated that the differences across year levels were statistically significant ($X^2 = 22.425$, df = 3, $p < 0.001$), suggesting that student feedback becomes progressively more positive with advancing year level. This pattern may reflect greater familiarity with courses, instructors, or university processes as students advance, highlighting the importance of considering year-level differences in evaluating student perceptions.

Classification of Sentiments into Positive and Negative

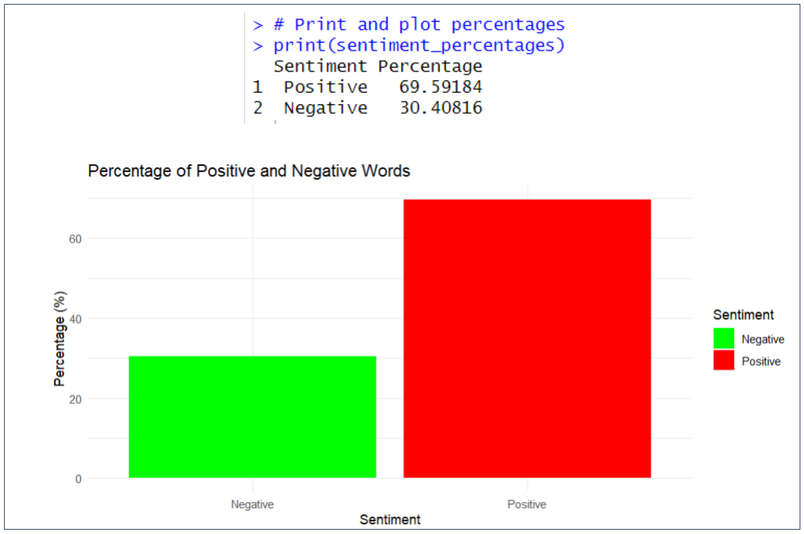

Figure 21. *Percentage of Positive and Negative Words*

Sentiment score analysis shows that the dataset has a mixed emotional tone, with a higher percentage of positive feelings than negative sentiments (Figure 21). Words that express praise, joy, and gratitude are frequently employed and represent positive attitudes. But there are also unpleasant feelings, especially when they are connected to challenges or failures. The bulk of the comments have a propensity toward confidence and contentment, and this balance between positive and negative attitudes offers insightful information about the overall sentiments and ideas of those who commented.

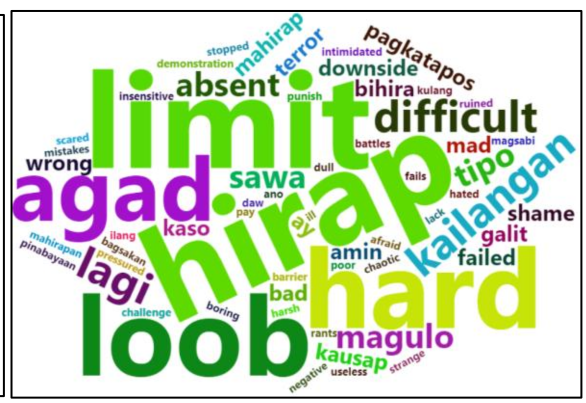

Figure 22. *Word Cloud of Positive Words* **Figure 23.** *Word Cloud of Negative Words*

The figures above (Figures 22 and 23) show the positive and negative word clouds extracted from the study. The positive word cloud, which is over half the size of the total word cloud, strongly emphasized the word *mabait* (kind), followed by *magaling* (intelligent), in the observation of the word cloud for positive and negative mood. Strongly favorable feedback was indicated by the magnitude of these two words, which were the most important. Numerous positive comments that expressed admiration for the instructor's attributes were almost the same size. The negative word cloud, on the other hand, had several terms with comparable proportions, making them easier to understand, like limit, *hirap* (difficult), hard, l*oob* (inside), and *agad* (quickly). Even while the negative adjectives were less common, they, however, highlighted important issues and problems.

The word clouds and bar plots provide clear analytical insights into student perceptions by visually emphasizing frequently mentioned positive and negative terms. Positive words such as *mabait* (kind), *magaling* (intelligent), and *masipag* (industrious) highlight teaching behaviors that students appreciate, including approachability, clarity in instruction, and dedication to student learning. Conversely, negative terms like *hirap* (difficult), hard, and pressure indicate challenges related to course pacing, workload, and instructional complexity. By linking these frequently used words to specific teaching strategies and classroom experiences, the visualizations allow faculty to identify strengths to reinforce and areas needing improvement, such as adjusting assignment difficulty, providing additional explanations, or pacing lessons more effectively. These figures, therefore, serve not only as summaries of sentiment but also as actionable guides for enhancing teaching methods and the overall learning environment.

Insights on Enhancement of Teaching Strategies

The analysis of student comments provided clear insights that can guide faculty in improving instructional strategies. Step 12 (word clouds) visually highlighted frequently mentioned positive and negative terms. Positive words such as *mabait* (kind), *magaling* (intelligent), love, and *masipag* (industrious) indicate that students appreciate approachable, dedicated, and effective teaching. Negative words like *hirap* (difficult), hard, pressure, and limit revealed areas where students face challenges related to course difficulty, workload, and time management. Step 13 (bar plots of top words) further quantified these trends, showing the most common positive and negative sentiments. Positive words were dominant, reflecting students' general satisfaction with teaching practices, while negative words highlighted recurring challenges that instructors can address. Step 14 (sentiment breakdown) summarized the overall proportion of positive and negative comments, providing a snapshot of the class-wide perception of teaching quality.

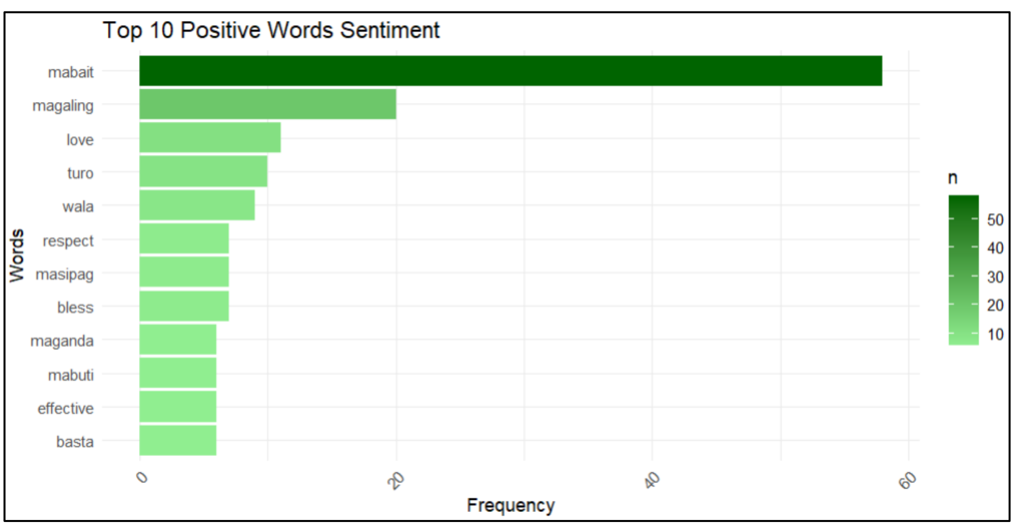

Figure 24. *Bar Plot of Top 10 Positive Words*

Based on the study, 69.59% of the general attitude conveyed is positive, suggesting that people have a generally positive opinion of the teachers. The most often used positive adjectives are *mabait* (kind), *magaling* (intelligent), and love, which express how much learners value the instructors' ability, friendliness, and enthusiasm for their work. Additional common positive terms like *turo* (teach), *masipag* (industrious), and effective highlight how much they value teachers who are committed, diligent, and successful in their strategies. Words like bless, respect, *basta* (just), and *maganda* (beautiful) demonstrate that pupils regard their instructors' interpersonal abilities and the classroom environment as a whole. The high frequency of these affirmative terms implies that students have a favorable opinion of their educational experiences, which shows that instructors who interact with students in a helpful and meaningful manner are well-liked.

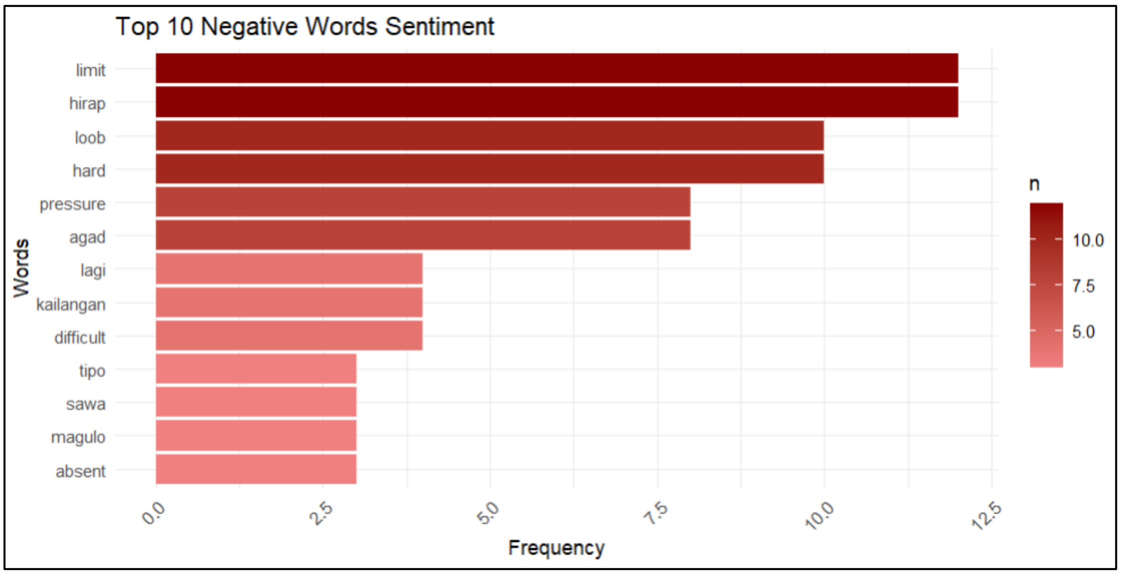

Figure 25. *Bar Plot of Top 10 Negative Words*

The bar plot of the top 10 negative terms is displayed in Figure 25. However, 30.41% of unfavorable opinions were also identified, which sheds light on areas that require improvement. Negative terms such as hard and *hirap* (difficulty) suggest that certain students find certain components of the coursework difficult. These opinions imply that some program components can be viewed as being overly challenging or demanding. Words such as pressure, limit, and *kailangan* (need) show that the student experiences stress as a result of the program's requirements. The results suggest that some students may be experiencing stress as a result of the workload or time management needs. There were words like *lagi* (always), *agad* (immediately), and *loob* (inner self) show that some students can feel pressured or under constant strain. These sentiments highlight how important it is to provide students enough time and assistance while still having a balance in the academic demand. Furthermore, the use of the term "absent" in the negative attitudes may be a reflection of worries about the attendance of the faculty member or the influence of outside variables such as academic or personal demands.

Through this analysis, a full representation of the student sentiment is offered by the mix of favorable and unfavorable opinions. Faculty may address areas where students struggle, including how to manage stress, workload, time, and difficulty levels, while also using student strengths, which include kindness, effectiveness, and dedication. Faculty members may improve a balanced and encouraging learning environment that encourages and motivates students to have a higher involvement and achievement by implementing changes in response to this input.

The analysis of student comments using NLP revealed both positive and negative sentiments that provide actionable insights for teaching and curriculum development. Positive terms indicate that students value instructors' approachability, dedication, and effective teaching strategies, suggesting that reinforcing these behaviors can enhance learning experiences. Negative terms highlight areas where students face challenges, pointing to potential adjustments in workload, pacing, and instructional clarity. Word clouds and bar plots visually summarized these trends, showing the frequency and emphasis of sentiment-bearing words, while also linking specific words to teaching behaviors and course components. Special attention was given to the bilingual nature of the dataset, with context-aware preprocessing and separate lexicons for Tagalog and English, ensuring that words with different meanings in each language were correctly interpreted, which enhanced the reliability of sentiment classification and allowed faculty to draw meaningful conclusions from students' feedback.

Conclusion and Future Works

This study analyzed student feedback from faculty evaluations in the Bachelor of Science in Computer Science program using natural language processing (NLP) techniques. The results revealed that most students expressed positive opinions about faculty approachability, teaching effectiveness, and engagement. Negative comments highlighted challenges such as course difficulty, workload, and classroom environment. Based on these findings, faculty can improve teaching strategies by maintaining supportive and approachable teaching while engaging students actively; adjusting course pacing, clarifying complex topics, and providing additional resources to reduce stress; and creating a welcoming classroom environment that encourages participation and collaboration. On the other side, program chairs and administrators can use these insights to guide curriculum improvements and faculty development programs.

The study has several limitations. First, it relied only on textual comments, which do not capture non-verbal cues such as facial expressions, tone of voice, or other environmental factors influencing student perception. Second, the dataset was limited to the BSCS program, so results could not be generalized to other fields. Third, the analysis did not integrate other forms of student feedback, such as surveys, class participation data, or peer evaluations, which could enrich the understanding of student experiences.

Future research can address these limitations by incorporating multiple feedback sources, including non-verbal cues, surveys, and peer reviews; applying advanced NLP techniques such as topic modeling or deep learning for more nuanced insights; and expanding the developed system to other programs or institutions, enabling school administrations to use it for faculty development programs and broader academic planning.

In conclusion, sentiment analysis of student feedback provides actionable guidance for faculty and administrators. By addressing identified challenges and reinforcing teaching strengths, instructors can enhance the learning experience, and instructional strategies can become more effective and responsive to student needs.

References

[1] Akhil, K., Vamsi, T. B., Soujanya, S., Harshitha, K. K. S., Mounisha, M., & Srihitha, N. (2024). Analyzing unstructured data: Natural language processing frameworks, challenges, approaches, and applications. In *2024 IEEE 4th International Conference on ICT in Business, Industry & Government (ICTBIG)* (pp. 1–9). IEEE. https://doi.org/10.1109/ICTBIG64922.2024.10911818

[2] Arifin, A., Suryaningsih, S., & Arifudin, O. (2024). The relationship between classroom environment, teacher professional development, and student academic performance in secondary education. *International Education Trend Issues, 2*(2), 151–159. https://doi.org/10.56442/ieti.v2i2.467

[3] Balahadia, F. F., & Comendador, B. E. V. (2016). Adoption of opinion mining in the faculty performance evaluation system by the students using Naive Bayes algorithm. *International Journal of Computer Theory and Engineering, 8*(3), 255–259. https://doi.org/10.7763/IJCTE.2016.V8.1054

[4] Balahadia, F. F., Fernando, M. C. G., & Juanatas, I. C. (2016). Teacher's performance evaluation tool using opinion mining with sentiment analysis. In *2016 IEEE Region 10 Symposium (TENSYMP)* (pp. 95–98). IEEE. https://doi.org/10.1109/TENCONSpring.2016.7519384

[5] Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student performance evaluation approaches in the higher education sector in a new era of ChatGPT: A case study. *Cogent Education, 10*(1), Article 2210461. https://doi.org/10.1080/2331186X.2023.2210461

[6] Chavan, R., Latthe, S., Dhorepati, M., Suryawanshi, A., Sharma, N., & Salge, A. (2024). Sentiment analysis using VADER and word cloud techniques. In *AIP Conference Proceedings* (Vol. 3217, No. 1, Article 020012). AIP Publishing. https://doi.org/10.1063/5.0234543

[7] Constantinou, C., & Wijnen-Meijer, M. (2022). Student evaluations of teaching and the development of a comprehensive measure of teaching effectiveness for medical schools. *BMC Medical Education, 22*(1), Article 113. https://doi.org/10.1186/s12909-022-03148-6

[8] Delgado, D. A., & Cabilles, R. (2024). Challenges in the implementation of online teaching and learning in Thailand: Insights for educational policy. *Isabela State University Linker: Journal of Education, Social Sciences and Allied Health, 1*(2), 54–65. https://doi.org/10.65141/jessah.v1i2.n6

[9] Deshpande, S. B., Tangod, K. K., Srinivasaiah, S. H., Alahmadi, A. A., Alwetaishi, M., Ong Michael, G. K., & Rajendran, S. (2025). Elevating educational insights: Sentiment analysis of faculty feedback using advanced machine learning models. *Advances in Continuous and Discrete Models, 2025*(1), Article 89. https://doi.org/10.1186/s13662-025-03933-9

[10] Dogra, V., Verma, S., Kavita, Chatterjee, P., Shafi, J., Choi, J., & Ijaz, M. F. (2022). A complete process of text classification system using state-of-the-art NLP models. *Computational Intelligence and Neuroscience, 2022*, Article 1883698. https://doi.org/10.1155/2022/1883698

[11] Facciolo, F., & Pittenger, A. (2024). A review of performance evaluation paradigms involving practice faculty. *American Journal of Pharmaceutical Education, 88*(11), Article 101293. https://doi.org/10.1016/j.ajpe.2024.101293

[12] Grimalt-Álvaro, C., & Usart, M. (2024). Sentiment analysis for formative assessment in higher education: A systematic literature review. *Journal of Computing in Higher Education, 36*(3), 647–682. https://doi.org/10.1007/s12528-023-09370-5

[13] Hasan, M., Ahmed, T., Islam, M. R., & Uddin, M. P. (2024). Leveraging textual information for social media news categorization and sentiment analysis. *PLOS ONE, 19*(7). https://doi.org/10.1371/journal.pone.0307027

[14] Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. F. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, *6,* Article 100059. https://doi.org/10.1016/j.nlp.2024.100059

[15] Kim, J., Li, X., & Bergin, C. (2024). Characteristics of effective feedback in teacher evaluation. *Educational Assessment, Evaluation and Accountability, 36*(2), 195–214. https://doi.org/10.1007/s11092-024-09434-9

[16] Lin, F., Li, C., Lim, R. W. Y., & Lee, Y. H. (2025). Empower instructors with actionable insights: Mining and visualizing student written feedback for instructors' reflection. *Computers and Education: Artificial Intelligence*, *8,* Article 100389. https://doi.org/10.1016/j.caeai.2025.100389

[17] Payandeh, A., Ghazanfarpour, M., Khoshkholgh, R., Malakoti, N., Afiat, M., & Shakeri, F. (2023). Views of students and faculty members on faculty evaluation by students: A systematic review. *Medical Education Bulletin, 4*(1), 659–671. https://DOI:10.22034/MEB.2023.386271.1073

[18] Poonputta, A., & Nuangchalerm, P. (2024). A model framework for enhancing twenty-first century competencies in primary school teachers within northeastern Thailand's sub-area. *International Journal of Learning, Teaching and Educational Research, 23*(1), 98–113. https://doi.org/10.26803/ijlter.23.1.6

[19] Raees, M., & Fazilat, S. (2024). Lexicon-based sentiment analysis on text polarities with evaluation of classification models. *arXiv*. https://doi.org/10.48550/arXiv.2409.12840

[20] Rongali, S. K. (2025). Natural language processing (NLP) in artificial intelligence. *World Journal of Advanced Research and Reviews, 25*(1), 1931–1935. https://doi.org/10.30574/wjarr.2025.25.1.0277

[21] Salgado, K. D., Arboleda, E., Cabardo, M. J., Obejera, C. M., & Dioses, J., Jr. (2024). Prediction model on the relationship of undergraduate grades and licensure examination performance of BS Agriculture and Biosystems Engineering. *Isabela State University Linker: Journal of Engineering, Computing and Technology, 1*(1), 15-32. https://doi.org/10.65141/ject.v1i1.n2

[22] Shahare, G., Manchalwar, J., Wankhede, K., Meshram, S., Kalbande, K., & Wyawahare, N. (2024). Exploring multidisciplinary data visualization and data analysis with R. In *2024 2nd International Conference on Advancements and Key Challenges in Green Energy and Computing (AKGEC)* (pp. 1–6). IEEE. https://doi.org/10.1109/AKGEC62572.2024.10869196

[23] Sharma, N. A., Ali, A. S., & Kabir, M. A. (2025). A review of sentiment analysis: Tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics, 19*, 351–388. https://doi.org/10.1007/s41060-024-00594-x

[24] Skeppstedt, M., Ahltorp, M., Kucher, K., & Lindström, M. (2024). From word clouds to Word Rain: Revisiting the classic word cloud to visualize climate change texts. *Information Visualization, 23*(3), 217–238. https://doi.org/10.1177/14738716241236188

[25] Sweta, S. (2024). *Sentiment analysis and its application in educational data mining*. Springer Nature. https://doi.org/10.1007/978-981-97-2474-1

[26] Takaki, P., & Dutra, M. L. (2023). Text mining applied to distance higher education: A systematic literature review. *Education and Information Technologies, 29*(9), 11241–11265. https://doi.org/10.1007/s10639-023-12235-0

[27] Yacoub, A. D., Slim, S., & Aboutabl, A. (2024). A survey of sentiment analysis and sarcasm detection: Challenges, techniques, and trends. *International Journal of Electrical and Computer Engineering Systems, 15*(1), 69–78. https://doi.org/10.32985/ijeces.15.1.7

Acknowledgement

I would like to express my gratitude to Dr. Charmaine G. San Jose, the College Dean, for granting the author permission to conduct this study. The researcher also appreciates the faculty and students of the BS Computer Science department for the comments and suggestions used as data in this research.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Artificial Intelligence (AI) Declaration Statement

The author acknowledges using ChatGPT, an artificial intelligence language model, to improve sentences, organize ideas, and enhance the clarity of the manuscript. The tool was not used for generating or analyzing data. All interpretations, findings, and conclusions in this study are the author's own. They have been carefully reviewed to ensure they are accurate, original, and follow ethical research standards.

Appendix

Table 7. Selected Tagalog Terms and English Translations

Tagalog Terms	English Translation	Context / Notes
mabait	kind	Frequently used positive descriptor of instructors' personality
magaling	excellent/great	Positive word describing teaching ability or skill
masipag	industrious/hardworking	Describes the dedication and effort of the instructor
mabuti	good	General positive descriptor for actions or character
maganda	good/beautiful	Positive comment about teaching or subject
hirap	difficult/hard	Negative sentiment about the difficulty of tasks or lessons
limit	limit/restriction	Refers to constraints in time, workload, or instructions
turo	teach / teaching	Positive or neutral word about teaching activity
loob	inner self/attitude	Context-dependent; can indicate internal motivation or disposition
agad	immediately/quickly	Negative or neutral; often refers to rushed tasks or deadlines
pressure	pressure/stress	Negative sentiment about stress caused by workload or requirements
wala	none/nothing	Negative sentiment indicating a lack of something (support, resources, clarity)
basta	just/simply	Neutral/positive; used in phrases like "basta okay" (just fine)
respect	respect	Positive, often about teacher-student interaction
bless	bless/God bless	Positive sentiment expressing gratitude or goodwill
favorite	favorite	Positive; indicates student preference or admiration