Understanding Student Sentiment Towards Informatics Engineering: Strategies to Attract High School and Vocational Graduates

Rahmat Taufik R. L. Bau[1*], Hermila A.[2], Mukhlisulfatih Latief[3]

[1,2,3]Department of Informatics Engineering, Faculty of Engineering, Universitas Negeri Gorontalo, Gorontalo City, Gorontalo, Indonesia
E-mail: [1*]rahmattaufik@ung.ac.id, [2]hermila@ung.ac.id, [3]mukhlis@ung.ac.id

(Received: 9 Dec 2024, revised: 15 Jan 2025, accepted: 17 Jan 2025)

Abstract

Higher education plays a crucial role in shaping the future of the younger generation, and in the ever-evolving digital era, technology has become an integral part of the education process. Amid the ongoing digital transformation, students' interest in the Informatics Engineering major is increasing; however, challenges remain in attracting high school (SMA) and vocational school (SMK) students to pursue this field. This research aims to provide a deeper understanding of students' sentiments toward the Informatics Engineering major and to formulate an effective promotional strategy to encourage high school and vocational school graduates to choose this path. To achieve these objectives, the research employs the TextBlob classification method, a natural language processing tool that assigns sentiment polarity scores (positive, neutral, or negative) to textual data. Sentiment analysis was conducted on responses collected through questionnaires, involving number of high school and vocational school students. The results of the sentiment analysis for high school (SMA) students reveal that out of 209 data points, 93 tweets (44.5%) were categorized as positive sentiment, citing career prospects and academic opportunities as key motivators. In contrast, For vocational school (SMK) students, among 135 data points analyzed, 50 tweets (37.0%) were categorized as positive sentiment, prioritizing practical skills and industry readiness. Based on the findings, the study formulates targeted promotional strategies. For SMA students, the focus should be on showcasing career prospects, technical skill development, and success stories in the tech industry. For SMK students, the promotion should emphasize practical, hands-on skills, industry partnerships, and job-readiness. This research provides recommendations for tailored promotional approaches to enhance students' awareness and interest in Informatics Engineering, thereby encouraging greater enrollment in the field.

Keywords: Students Sentiment Analysis, Informatics Engineering, Promotion Strategies, Textblob.

I. INTRODUCTION

Higher education is an important milestone in shaping the future of the younger generation [1], and in the ever-evolving digital age, the role of technology is increasingly becoming an integral part of the education process [2]. Young people today are growing up in an environment dominated by information and communication technology (ICT), and skills in using technology are key to future success [3].

From an early age, young people are exposed to various technological devices, ranging from smartphones to computers. This creates a generation that is familiar with technology and has the ability to access information quickly and efficiently. In the context of education, technology has changed the way we learn and teach. Online learning platforms, educational apps and other digital resources allow students to learn independently and collaboratively, even outside the classroom.

However, the impact of technology is not just limited to the way we learn. Changes in student behavior and preferences are also reflected in their career choices [4]. In the midst of the ongoing digital transformation, students' interest in Informatics Engineering is increasing. The major offers broad and exciting career opportunities in various sectors, ranging from information technology, cyber security, to software development. However, challenges still exist in attracting high school and vocational school students to major in Informatics Engineering.

In the context of higher education, the comparison between high school and vocational students is an interesting subject of study. Research shows that vocational students (SMK) often have lower levels of access and perseverance than academic students (SMA) in achieving higher education [5]. This could be due to a variety of factors, including different curricula and differences in the emphasis on academic education versus practical skills in high school.

However, interestingly, studies have also shown that students who continue their education in the same field as their vocational education background (SMK) perform better in higher education compared to those with an academic background (SMA) [6]. Although they may initially face challenges in adjusting to the academic environment, students who graduate from SMK often show great perseverance and dedication, as well as strong practical skills that they gained from their vocational education [7].

In addition, research also shows that vocational education can have a positive impact on graduation rates and earnings later in life. Although few choose to go on to higher education after graduating from SMK, those who do tend to have higher graduation rates and higher earnings in their careers [8]. This suggests that vocational education can be a solid foundation for future success, regardless of whether students choose to continue formal education after graduating from high school or not.

Thus, in the context of this study, the comparison between high school and vocational school graduates is important to consider in analyzing their sentiments towards Informatics Engineering majors. With a better understanding of the differences in students' educational backgrounds and experiences, researchers can identify factors that might influence their preferences towards this major and formulate appropriate promotional strategies to increase their participation in higher education in Informatics Engineering.

This study aims to identify sentiment patterns that may exist among students, as well as investigate the relationship between their educational background and their perception of this major. Furthermore, this research also serves as a foundation for designing appropriate promotional strategies to increase students' awareness and interest in the major. By understanding students' needs and preferences, universities can develop more effective and relevant promotional programs, which can provide first-hand insights into career opportunities in Informatics Engineering.

As a result, it is hoped that this research can make a valuable contribution in overcoming the challenges in promoting Informatics Engineering majors and increasing the interest and participation of high school and vocational school students in this field.

II. LITERATURE REVIEW

A. Sentiment Analysis

Sentiment analysis has become an essential component of natural language processing (NLP), providing valuable insights into subjective information within text. Recent advancements in deep learning, particularly with models like BERT (Bidirectional Encoder Representations from Transformers), have significantly improved the precision and efficiency of sentiment analysis.

In Bau’s study [9], he evaluated the use of a summarization-based sentiment analysis framework on online discourses. By integrating summarization, tokenization, and sentiment classification using BERT, the study achieved an impressive 86% precision rate. This research is particularly relevant for exploring public discourse sentiment, as it showcases the ability of sentiment analysis to handle large-scale, diverse textual data while maintaining high accuracy. Such methodologies can also be applied to analyze student feedback and perceptions, making it an essential reference for studies focusing on educational sentiment analysis.

Another noteworthy contribution, highlights the application of sentiment analysis in the education domain. By investigating students’ sentiments toward ChatGPT, Bau demonstrated how advanced sentiment analysis tools can provide insights into students’ acceptance and perceptions of AI-driven educational technologies [10]. This aligns with the journal's focus on understanding and addressing factors that influence students’ engagement with technology in education, providing a robust framework for analyzing their preferences and concerns.

These studies emphasize the importance of adopting advanced sentiment analysis techniques in various contexts. Bau’s work exemplifies how sentiment analysis can be leveraged to extract nuanced insights from complex textual data, whether in online discourse or educational settings. By relating these findings to the journal’s themes, future research can further explore the intersection of NLP and education to improve student engagement and program promotion strategies.

B. Textblob

Sentiment analysis, a pivotal aspect of natural language processing (NLP), involves determining the emotional tone behind a body of text to understand the sentiments expressed. TextBlob, a Python library, offers a straightforward API for performing common NLP tasks, including sentiment analysis. Compared to more advanced methods like BERT (Bidirectional Encoder Representations from Transformers) or VADER (Valence Aware Dictionary and sEntiment Reasoner), TextBlob stands out for its simplicity and ease of use, making it an excellent choice for quick analyses and small-scale projects. While BERT leverages deep learning to provide highly accurate sentiment predictions with contextual understanding, and VADER is optimized for social media text with its rule-based approach, TextBlob is less computationally intensive and accessible. However, its reliance on pre-defined lexicons and lack of contextual awareness can limit its accuracy in more complex datasets.

Hazarika [11] applied TextBlob to analyze sentiments in Twitter data. They demonstrated how TextBlob can classify sentiments into positive, negative, or neutral categories, providing valuable insights into public opinions expressed on social media platforms.

Similarly, the paper by Gujjar [12] explored the use of TextBlob in analyzing customer feedback for market surveys. The study highlighted TextBlob's utility in understanding

DOI: 10.34148/teknika.v14i1.1165
ISSN 2549-8037, EISSN 2549-8045

customer sentiments, thereby aiding decision-making processes in business contexts.

These studies underscore TextBlob's versatility and effectiveness in sentiment analysis across various domains, including social media monitoring and market research. Its ease of use and reliable performance make it a valuable tool for extracting and interpreting sentiments from textual data.

C. State of the Art

Research on students' interest in pursuing higher education in Informatics Engineering programs has explored various dimensions, including curriculum influence, interest segmentation, and differences in academic backgrounds. Kamanjaya investigated the impact of the 2013 Curriculum on SMA and SMK students' interest in studying Informatics Engineering Education [13]. Their findings revealed a significant correlation between curriculum implementation and students' interest, emphasizing the role of educational policies in shaping preferences. However, their focus was on policy analysis, whereas the proposed study will explore students' sentiments using TextBlob sentiment classification, offering a more perception-driven perspective.

Suprapty [14] employed Fuzzy C-Means clustering to classify students' interest in the Informatics Engineering program at Samarinda State Polytechnic. Their research identified clusters with varying levels of interest, providing a quantitative understanding of student preferences. While effective in segmenting interest, this method does not address emotional or perception-based factors. The proposed study aims to fill this gap by utilizing TextBlob sentiment analysis to uncover nuanced insights into students' attitudes and perceptions toward Informatics Engineering.

Meanwhile, Nugroho analyzed differences in academic performance between SMA and SMK graduates in Electrical Engineering at Surabaya State University [15], finding that SMA graduates generally achieved better academic outcomes. Although conducted in a different academic context, their findings highlight the significance of educational background in shaping academic preferences and performance. The proposed research builds on this understanding by examining how SMA and SMK graduates' backgrounds influence their sentiments toward Informatics Engineering. By focusing on sentiment analysis, this study aims to provide actionable insights to inform effective promotional strategies for increasing interest in Informatics Engineering programs.

III. RESEARCH METHOD

This research will be carried out in several stages to ensure a comprehensive and structured approach to understanding students' sentiments toward the Informatics Engineering program. The following steps outline the methodology:

1. **Data Collection**: The first step involves gathering data through questionnaires to collect sentiment information from high school (SMA) and vocational school (SMK) students regarding their perceptions of the Informatics Engineering major. The expected outcome at this stage is a complete and representative dataset that captures a wide range of student sentiments from both educational backgrounds. Success will be measured by the response rate and the adequacy of the collected data.
2. **Data Preprocessing**: The collected data will undergo preprocessing, including data cleaning, tokenization, stopword removal, and stemming. The goal of this stage is to prepare the textual data for input into the TextBlob classification model. The expected outcome is clean data that is ready for sentiment analysis. Success will be determined by the ability to remove noise and errors from the data, and by the generation of an accurate and structured textual data representation.
3. **Data Analysis using TextBlob**: In this stage, the prepared dataset will be used to classify students' sentiments using the TextBlob library, which employs a rule-based algorithm to analyze the polarity of textual data, assigning scores from -1 (negative) to +1 (positive), with values near zero considered neutral. Sentiment classification accuracy will be assessed by comparing the model's predictions against a manually labeled dataset serving as the ground truth. Accuracy will be calculated as the percentage of correctly classified sentiments, while additional metrics such as precision, recall, and F1-score will provide a more comprehensive evaluation of the model's performance. The expected outcome is a reliable TextBlob classification model capable of accurately analyzing students' sentiments, forming the basis for further insights and promotional strategies.
4. **Data Comparison**: The final step involves comparing the sentiment classifications between SMA and SMK graduates. This analysis will compare the sentiment results between the two groups of students based on the previously prepared data. The expected outcome is the identification of significant differences in sentiments between SMA and SMK students towards the Informatics Engineering major. Success will be determined by the presence of a noticeable difference in the distribution of sentiments between the two student groups.

These stages will provide a thorough understanding of the factors influencing students' perceptions of the Informatics Engineering major and help in formulating strategies to promote this field of study effectively.

IV. RESULT AND DISCUSSION

A. Data Collection

The results of this study show that we managed to collect data from 377 respondents spread across several high schools and vocational schools in Gorontalo City. Data collection was conducted directly in the field using stratified random sampling. At the time of data collection, respondents were asked to fill out a questionnaire that included information about their school of origin, gender, current major, and their consideration for continuing their studies to a higher level.

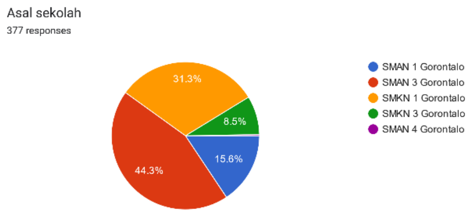

Figure 1. Respondent's School of Origin

The results of the data collection in Figure 1 provide an initial picture relevant to the research on the comparison of high school and vocational school students' sentiments towards Informatics Engineering majors in pursuing higher education. Most of the respondents came from high schools, specifically SMA Negeri 3 Gorontalo (44.3%) and SMA Negeri 1 Gorontalo (15.6%), while the remaining 39.8% came from SMK. This composition shows that the majority of respondents have a general education background compared to vocational education.

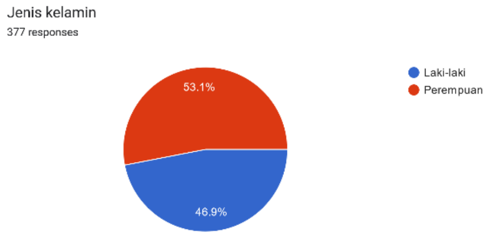

Figure 2. Gender of Respondents

Based on gender (Figure 2), the majority of respondents were female (53.1%) compared to male (46.9%), providing an opportunity to see variations in sentiment based on gender in the context of choosing an Informatics Engineering major. In addition, when asked about plans to continue higher education, 232 out of 377 respondents expressed a desire to continue, 112 respondents were still considering by selecting the option "maybe," and 33 respondents expressed no interest.

This data provides a strong basis to further explore how educational backgrounds and perceptions of higher education, particularly in Informatics Engineering, differ between SMA and SMK students. To administer this, we distributed a questionnaire to the students, asking, *"Apa menurut Anda memilih Teknik Informatika sebagai jurusan Anda saat kuliah?"* (What are your thoughts on choosing Informatics Engineering as your major in college?). This approach helped gather qualitative insights into their motivations, concerns, and expectations. Understanding these patterns can significantly influence strategies aimed at guiding students' decisions regarding further studies and promoting the Informatics Engineering major effectively.

B. Data Preprocessing

Data preprocessing is a crucial step in text analysis to ensure that the data used in the research is clean and well-structured. The first step is to remove duplicate data. Duplication in the dataset can lead to biased analysis results, as repeated entries may give more weight to some responses, ignoring the diversity of opinions present. Using the pandas library, we removed duplicates from the dataset using the drop_duplicates() function, which automatically identifies and removes rows that have the same value. This ensures that the dataset used for analysis does not contain redundant data, providing more representative results.

After that, the next step is the cleaning of the text from irrelevant symbols and numbers. In text analysis, symbols or numbers that are not related to the research topic can interfere with the analysis process. Therefore, symbols, numbers, and other characters that do not have important meaning in the context of the research are removed. This process is done using regular expressions, which allow matching of certain character patterns in the text. With the help of the re.sub() function, we were able to easily remove numbers and symbols, ensuring only relevant text remained in the dataset.

After removing duplicate data and irrelevant symbols, the next step is to remove empty data. Empty or null data can reduce the quality of the analysis, as it can cause errors or inaccuracies in the results. Therefore, all rows containing empty values were removed using the dropna() function from pandas. This process ensures that the dataset used contains only complete data, improving the accuracy of the analysis to be performed.

After the data was cleaned, the next step was to convert all characters in the text to lowercase. This aims to equalize the writing of words that may have different capitalization, but the same meaning. For example, the words "Technology" and "technology" should be considered identical in sentiment analysis. For that, we use the .str.lower() method to convert the entire text to lowercase, simplifying the text to be analyzed and reducing possible differences that are only caused by capitalization.

The next process is word normalization. In this stage, variations in the writing of words that have similar meanings are equalized. For example, the word "gak" which is a variation of "not", or "bgt" which means "really", is replaced with its raw word. This normalization is done by replacing certain words using a pre-prepared normalization dictionary. This helps to ensure consistency in the dataset and avoid analyzing the same words but written differently.

Then, the stop words removal process is performed. Stop words are common words such as "at", "which", and "with" that do not contribute significantly to sentiment analysis. To remove stop words, we use the Sastrawi library, which provides a list of Indonesian stop words. Sastrawi's remove() function was used to remove these words, so that the analysis could focus on words that are more meaningful in the context of the study.

After that, the cleaned data went through tokenization. Tokenization breaks the text into small units, usually words, which can then be analyzed. Using the nltk library, we broke the text into individual tokens using the word_tokenize() function. Tokenization allows analysis to be done at the word level, which is very important in sentiment research.

Finally, a stemming process is performed to convert the affixed words into their base form. For example, the word "run" is converted into "run", or "eat" is converted into "eat".

ISSN 2549-8037, EISSN 2549-8045
DOI: 10.34148/teknika.v14i1.1165

This stemming process is done with the help of the Sastrawi library, which simplifies words so that the analysis focuses more on the basic meaning of the word. With all these steps, the previously raw data is now ready for further analysis, providing more accurate and reliable results for sentiment research related to high school and vocational school students' attitudes towards Informatics Engineering majors.

C. Data Analysis

In analyzing sentiment data we use Python. The data analyzed comes from the "Interest" column which contains texts from the questionnaire, which are then processed to classify each opinion based on positive, neutral, or negative sentiment.

The process starts by retrieving data from the "Interest" column in the dataset and storing it in the data_opinion variable. Then, the polarity variable is initialized with a value of 0, which will be used to calculate the total number of sentiment polarities in the data. Next, a status list is initialized to store the sentiment status (Positive, Neutral, or Negative) for each tweet, whose length is equal to the number of tweets. Initially, each tweet in the status list is assigned the value 'None'.

Using the for loop, each opinion in data_opinion is analyzed one by one. Before performing the analysis, the code ensures that the element being analyzed is a string and not a NaN (empty) value. The analysis process is performed using the TextBlob library, which provides a method to calculate sentiment polarity. This polarity is a value between -1 and 1, indicating the degree of negativity (-1) or positivity (1) of a text, with a value of 0 indicating neutrality.

After analysis, the polarity value of each opinion is added to the polarity variable, which will give an idea of the general tendency of the entire dataset. Based on the resulting polarity value, the opinion is categorized into one of three sentiment categories:

- Positive if the polarity is greater than 0.
- Neutral if the polarity is equal to 0.
- Negative if the polarity is smaller than 0.

The number of opinions in each of these categories (positive, neutral, negative) is then counted and stored in the variables total_positive, total_neutral, and total_negative. For each opinion analyzed, its sentiment status is also stored in the status list in the index corresponding to that opinion.

Once all opinions have been analyzed, the analysis results are printed to show the number of opinions in each category (positive, neutral, and negative) as well as the total number of opinions analyzed. This process provides a clear picture of how the respondents feel about the topic, whether they are more likely to have a positive, neutral, or negative view in relation to the major under study.

This code allows researchers to conduct efficient sentiment analysis and provide results that can be used to explore a deeper understanding of the opinions of high school and vocational school students towards Informatics Engineering majors, which will help in designing educational policies or programs that are more in line with their interests and needs.

Based on the analysis above, here are the results we get from processing the opinions of high school students and vocational students as shown in Table 1.

Table 1. Sample Results of High School Students' Sentiments

vel_0	index	*Ketertarikan*	Interest	Classification
0	0	*kurang tarik kurang sesuai paket mata ajar*	not attractive enough not in accordance with the teaching package	Negative
1	1	*jurus perlu*	necessary majors	Neutral
2	2	*memang tarik*	it's interesting	Positive
3	3	*kurang tarik*	less attractive	Positive
4	4	*seru*	exciting	Positive
5	5	*peluang kerja banyak*	lots of job opportunities	Neutral
6	6	*tarik*	interest	Neutral
7	7	*teknik informatika bukan ahli minat*	informatics engineering is not an expert of interest	Neutral
8	8	*saya tarik*	I interested it	Positive
9	9	*tarik informatika peluang kerja nya banyak*	There are lots of job opportunities in informatics	Neutral
10	10	*tarik milik bakat informatika*	interest belongs to informatics talent	Neutral
11	11	*karna bukan minat*	because it's not of interest	Neutral

Based on the results of sentiment analysis of 209 data obtained, there are 93 tweets categorized as positive sentiment, 88 tweets with neutral sentiment, and 28 tweets with negative sentiment. This indicates that most respondents showed a positive interest or view towards the topic analyzed, with a significant proportion also showing a neutral view. Only a small proportion of respondents gave a negative response to the topic, which may indicate disinterest or challenges in understanding the material or subject matter discussed.

After the analysis process, the accuracy of the model used to classify the sentiment of the tweets was 0.76, which means that the model managed to correctly classify the sentiment of 76% of the total tweets analyzed. Although this accuracy shows good performance, further evaluation by looking at the classification report provides a more detailed picture of the model's performance in each sentiment category.

In the classification report, for the Negative category, the model obtained a very low precision and recall (0.00), which indicates that the model struggled to detect negative sentiments, possibly due to the limited number of tweets with negative sentiments in the dataset. This causes the model to be ineffective in classifying tweets as negative sentiment even though there are some tweets with such sentiments.

For the Neutral category, the model's precision is 0.88, which indicates that the model is quite good at identifying truly neutral tweets. However, the recall was slightly lower (0.78), indicating that some tweets that should have been neutral were not perfectly detected. Nonetheless, overall, the model is able to recognize neutral sentiments quite well.

As for the Positive category, the model performed quite well with a precision of 0.72 and a very high recall of 0.95. This means that the model is very effective in identifying positive tweets, although there is a slight error in classifying some positive tweets. This result shows that the model is more likely to accurately classify tweets with positive sentiments.

Overall, although the overall accuracy of the model is 076 or 76%, the results of this analysis show that the model performs best in identifying positive sentiments, while the negative category requires further improvement.

Table 2. Sample Results of Vocational School Students' Sentiments

Index	Unnamed:	*Ketertarikan*	Interest	Classification
0	0	*minat pribadi*	personal interest	Neutral
1	1	*tau*	you know	Neutral
2	2	*tarik buat program game*	drag it to create a game program	Negative
3	3	*suka pemograman sejak kecil*	I like programming since childhood	Neutral
4	4	*yaaa kurang tarik aja sih*	yeah it's just not interesting enough	Negative
5	5	*kembang teknologi informasi komunikasi*	development of information communication technology	Neutral
6	6	*tarik bisa buat game*	pull it can make a game	Negative
7	7	*buat tarik saat dunia pusat ada teknologi nah lulus teknik informatika sangat peran penting hal*	To make it interesting when the world is at the center of technology graduating informatics engineering has a very important role	Positive
8	8	*sangat tarik*	very attractive	Positive
9	9	*tarik teknik informatika minat teknologi pemrograman peluang karier luas dunia digital tantang dalam ajar konsepkonsep teknis kompleks jadi yang tantang*	interest in informatics engineering interest in programming technology broad career opportunities in the digital world challenges in teaching complex technical concepts so it's challenging	Positive
10	10	*lebih mudah dapat kerja*	easier to get work	Neutral

Based on the results of sentiment analysis of SMK students' opinions (Table 2), it can be seen that the distribution of sentiments is slightly different compared to the previous high school students. Of the 135 data analyzed, 50 tweets are categorized as positive sentiment, 47 tweets as neutral, and 38 tweets as negative.

Unlike the results of the analysis of high school students' opinions, which showed a stronger positive tendency, the sentiment analysis of vocational students showed a more balanced distribution, although there was still a tendency to give positive sentiments. One reason that could explain this difference is the difference in educational context between high school and vocational students, where vocational students may be more pragmatic in considering the relevance of their chosen major to the world of work, while high school students tend to be more idealistic in assessing their educational choices.

In terms of the accuracy of the model used for sentiment analysis, the measurement results show that the overall accuracy is 0.69, which means that the model can correctly classify tweets for 69% of the total data analyzed. Based on the classification report, for the Negative category, the precision reaches 0.67, which indicates that the model is quite good at recognizing tweets with negative sentiments, but the recall is only 0.25, which indicates that the model is not very successful in detecting all tweets with negative sentiments. This may be due to the limited amount of negative data in the dataset. For the Neutral category, the model's precision and recall are quite good, at 0.89 and 0.80 respectively, which indicates that the model is able to identify neutral tweets very well.

The Positive category shows a precision of 0.56, which is lower than the other categories, although the recall is quite high at 0.90. This shows that the model was more successful in finding positive tweets, although there were some errors in identifying tweets that were actually positive. In addition, the None category had perfect precision and recall (1.00), even though there was only one data point, indicating that the model was able to very well identify tweets that did not have a clear or relevant sentiment.

Overall, these results show that the model still needs improvement, especially in improving its ability to identify negative sentiments more effectively. However, the model did quite well in identifying neutral and positive sentiments, with an overall accuracy of 69%.

D. Sentiment Comparison

SMA curricula are generally academically oriented, preparing students for higher education and fostering an appreciation for theoretical fields like Informatics Engineering. This academic environment could explain the higher positive sentiments among SMA students, as they perceive the major as a pathway to promising career prospects in the tech industry.

Conversely, SMK programs emphasize practical skills and direct workforce entry, aligning with students' expectations for immediate job readiness. The more theoretical nature of Informatics Engineering may not resonate with their desire for hands-on training, leading to increased neutral or negative sentiments. Additionally, some SMK students may feel that their current technical skills suffice for employment, perceiving further academic study as unnecessary.

These findings align with previous studies indicating that vocational students often prioritize immediate employment and practical skill application over extended academic pursuits. For instance, research has shown that vocational students are more inclined to enter the workforce directly after high school, valuing hands-on experience and job-specific training [16]

The word frequency analysis results for each school category (high school and vocational school) provide a deeper insight into the differences and similarities in students' thoughts and perspectives regarding the major they are considering. From the analysis, it can be seen that some words appear in both categories, but there are also differences in the order and frequency of the words that emerge.

For high school students, the most frequently occurring word is “(ter) tarik” (attracted), with 158 occurrences, reflecting their interest or attraction to something, followed by the word “informatika” (informatics) with 100 occurrences (Figure 3). This indicates that many high school students are interested in the field of informatics engineering. Words like “teknik” (engineering) with 59 occurrences, “jurus (an)” (major) with 51 occurrences, and “kerja” (work) with 35 occurrences also reflect their focus on career opportunities and their choice of major. Additionally, words indicating evaluation or feelings, such as “kurang” (lacking) with 24 occurrences, also appear, which may suggest that some high school students feel there is a deficiency or mismatch in the major they are interested in.

Figure 3. High School Students' Sentiment Trends

On the other hand, for vocational school students, the most frequent word is “(ter) tarik” (attracted) with 75 occurrences, which shows that although this word appears less often than in the high school category, attraction remains a dominant theme (Figure 4). Other top words in the vocational school students' list include “jurus (an)” (major) with 32 occurrences, reflecting their focus on major choices, and “informatika” (informatics) with 27 occurrences, though the frequency is lower compared to high school students. Words like “tidak” (not) with 24 occurrences, “teknik” (engineering) with 16 occurrences, and “minat” (interest) with 14 occurrences also appear, which may indicate a more skeptical or less enthusiastic view towards certain majors, particularly informatics engineering. Furthermore, the word “kerja” (work) with 13 occurrences suggests that vocational students are more focused on job opportunities, aligning with their more practical orientation towards entering the industrial workforce.

Figura 4. Vocational School Students' Sentiment

When linking the word frequency analysis with the sentiment differences expressed by high school and vocational school students regarding pursuing higher education in informatics engineering, high school students tend to show more positive sentiment, as reflected in the high frequency of words like “(ter) tarik” and “informatika,” which indicate their interest in the field. Words such as “kerja” and “teknik” also suggest that they are considering future career opportunities. This positive sentiment is consistent with their tendency to be more open to continuing their studies at a higher level, as reflected in the data, where most high school students expressed interest in pursuing higher education.

In contrast, vocational school students, although using similar words such as "kerja" and "teknik," show a slightly more neutral or even negative sentiment, as seen from the frequency of words like “tidak” (not) and “minat” (interest) in their analysis. This may reflect their uncertainty or lack of interest in higher education in informatics engineering, which is perceived as more theoretical and academic compared to the more practical programs they expect. Therefore, while words related to engineering and work also appear in vocational school students' responses, their focus on direct entry into the workforce and practical skills may influence their more neutral or negative sentiment toward higher education in this field.

Thus, the differences in word frequency used by high school and vocational school students provide a clearer picture of the factors influencing their varying sentiments toward the informatics engineering major. High school students are more likely to see this major as an attractive and beneficial pathway for higher education, while vocational school students are more concerned about whether higher education in this field aligns with their career needs, leading to more neutral or even negative sentiments.

E. Recommendations

Based on the sentiment analysis, word cloud, and demographic data of the respondents, several strategies can be implemented to effectively promote the Informatics Engineering major to both high school (SMA) and vocational school (SMK) students, boosting awareness and interest in this field.

For high school students (SMA), the promotion should primarily focus on highlighting career prospects and job opportunities. Sentiment analysis reveals that most high school students have a positive view of the Informatics Engineering major, especially regarding its potential for employment. Therefore, promotional strategies should emphasize the broad job opportunities available in various technology industries. Words like “kerja” (work) and “informatika” (informatics), which appear frequently in the word cloud, should be incorporated to underline the promising career paths in digital and technological sectors. Additionally, the promotion should highlight the development of relevant technical skills, such as application development and cybersecurity, which are highly valued in today's job market. By showcasing success stories from alumni who have worked in the technology industry or started their own digital ventures, the promotion can attract high school students interested in building a successful career. Furthermore, the introduction of a curriculum that not only provides theoretical knowledge but also emphasizes practical experience, such as hands-on projects or internships, can significantly enhance the appeal of this major. Engaging visual content, especially through social media and promotional videos, can be used to further captivate the attention of high school students.

For vocational school students (SMK), the promotional strategy should be more focused on the practical applications of the Informatics Engineering field. Sentiment analysis and word cloud results suggest that SMK students show more neutral or negative sentiment towards the major, with terms like “tidak” (no) and “minat” (interest) appearing more frequently. To address this, promotional messages should emphasize the immediate, practical skills that students can apply directly in the workforce, such as software development, application creation, and network system management. The promotion should stress that an education in Informatics Engineering equips students with the technical expertise needed for hands-on roles in various industries, providing them with skills that are immediately applicable in the job market. Additionally, it’s essential to link higher education in this field to career opportunities, emphasizing that the major not only prepares students for academic roles but also for direct entry into professional positions, such as software development, network administration, or database management. Highlighting internship programs and partnerships with industry leaders can also be an effective way to appeal to SMK students. To further engage this group, showcasing success stories from SMK alumni who pursued Informatics Engineering and achieved career success, either by working with top companies or starting their own technology businesses, can provide concrete examples of the career potential in this field.

Furthermore, it's crucial to recognize the differences in language use and sentiment between the two groups. Promotional messages for high school students should emphasize long-term career opportunities, utilizing terms like "kerja" (work) that frequently appear in the word cloud, while promotions for vocational students should be centered around the practical applications of skills they can directly use in the workforce, focusing on terms like "tidak" (no) to address their concerns or lack of interest. Tailoring the message to each group’s specific needs and language will make the promotion more impactful.

Lastly, the choice of promotional media should align with the preferences of these students. Since both high school and vocational students are highly active on social media platforms, the promotion should focus on platforms such as Instagram, YouTube, and TikTok. Short, engaging videos that showcase the exciting aspects of learning Informatics, along with inspiring talks from IT professionals or practical tutorials on application development, will resonate well with these students. Additionally, data-driven campaigns using visuals like infographics, which display career prospects in the Informatics Engineering field or statistics on the demand for technology professionals, can be highly effective in catching students' attention.

By adopting these tailored strategies, the promotion of the Informatics Engineering major can be more effective in attracting high school and vocational school students, increasing their interest in pursuing higher education in this dynamic field.

V. CONCLUSION

The analysis of sentiment, word cloud data, and demographic information reveals a clear distinction between the perspectives of high school (SMA) and vocational school (SMK) students regarding the Informatics Engineering major. High school students exhibit a predominantly positive sentiment, viewing the major as a gateway to promising career opportunities and higher education. Their enthusiasm reflects their aspirations for academic achievement and long-term success in the tech industry. On the other hand, vocational school students demonstrate more neutral or skeptical sentiments, focusing on the practical application of skills and immediate job-readiness, often prioritizing hands-on experiences over theoretical pursuits.

To bridge these differences and effectively engage both groups, tailored promotional strategies are key. For high school students, the focus should be on highlighting the expansive career prospects and the transformative impact of Informatics Engineering on the digital world. Sharing inspiring success stories, showcasing advanced technical skill development, and emphasizing hands-on learning opportunities can resonate with their ambitions. Meanwhile, for vocational school students, promotions should underscore the immediate relevance of Informatics Engineering to real-world challenges. Demonstrating how the major equips them with industry-specific skills and enhances their employability will address their practical concerns and build trust in the academic pathway.

By combining these targeted approaches with innovative outreach methods, the Informatics Engineering major can attract diverse students and spark their interest. Leveraging social media platforms and creating engaging, data-driven content tailored to the preferences of each group will amplify the message and connect with them more effectively.

With these efforts, Informatics Engineering can position itself as a compelling choice for students, paving the way for a brighter future in the digital era. By inspiring both high school and vocational school students to pursue higher education in this dynamic field, we can nurture a new generation of tech-savvy professionals ready to shape the future.

ACKNOWLEDGEMENT

This research was funded by Universitas Negeri Gorontalo through Riset Akselerasi Kolaborasi Perguruan Tinggi (RAKPT) scheme Riset Kerjasama Fakultas (RKF) 2024 via Decree of Rector of Universitas Negeri Gorontalo No: 798/UN47/HK.02/2024.

REFERENCES

[1] A. Anwari, "Peningkatan Mutu Sistem Pendidikan Dengan Cara Meningkatkan Hal Yang Menunjang Kebijakan Sistem Zonasi," 2019.

[2] N. A. Ahmad, "Mengintegrasi teknologi komunikasi dan maklumat sebagai media pengajaran bahasa dalam kalangan kanak-kanak awal umur," *Jurnal Pendidikan Bahasa Melayu*, vol. 4, no. 2, pp. 36–43, 2016.

[3] V. Fransisca and W. E. Saputro, "Leaders Shaping The Future: Required Management Skills," *American Journal of Economic and Management Business (AJEMB)*, vol. 2, no. 6, pp. 226–230, 2023.

[4] J. E. Dangoy and D. V Madrigal, "Career Preferences and Factors Influencing the Career Choice of Senior High Students of a Catholic School," *Philippine Social Science Journal*, vol. 3, no. 2, pp. 95–96, 2020.

[5] M. Farías and M. P. Sevilla, "Effectiveness of vocational high schools in students' access to and persistence in postsecondary vocational education," *Res High Educ*, vol. 56, pp. 693–718, 2015.

[6] H. Wen-min, "Choosing General Senior High School or Vocational Senior High School——Survey on Further Education and Employment Aspirations of Junior High School Graduates," *Vocational and Technical Education*, 2013.

[7] O. Romanova, "Embedding employability skills into vocational education and training: What works best for students' self-evaluation and aspirations?," *Journal of Teaching and Learning for Graduate Employability*, vol. 13, no. 1, pp. 20–36, 2022.

[8] J. H. Bishop and F. Mane, "The impacts of career-technical education on high school labor market success," *Econ Educ Rev*, vol. 23, no. 4, pp. 381–402, 2004.

[9] R. T. R. L. Bau, "Summarizer Precision Value on Tribunnews Gorontalo in the Implementation of Online Discourse Sentiment Analysis," *Sinkron: jurnal dan penelitian teknik informatika*, vol. 8, no. 4, pp. 2422–2428, 2024.

[10] R. T. R. L. Bau, A. Hermila, I. Farman, L. M. Hidayat, and S. Salim, "AI Perspectives in Education: A BERT-based Exploration of Informatics Students' Attitudes to ChatGPT," in *2023 9th International Conference on Education and Technology (ICET)*, 2023, pp. 36–41. doi: 10.1109/ICET59790.2023.10435040.

[11] D. Hazarika, G. Konwar, S. Deb, and D. J. Bora, "Sentiment Analysis on Twitter by Using TextBlob for Natural Language Processing.," *ICRMAT*, vol. 24, pp. 63–67, 2020.

[12] J. P. Gujjar and H. P. Kumar, "Sentiment analysis: Textblob for decision making," *Int. J. Sci. Res. Eng. Trends*, vol. 7, no. 2, pp. 1097–1099, 2021.

[13] I. M. A. Kamanjaya, K. Agustini, M. Si, I. G. M. Darmawiguna, and S. Kom, "Pengaruh Penerapan Kurikulum 2013 dan Penerimaan Mahasiswa Baru Terhadap Minat Studi Lanjut Siswa SMA dan SMK di Jurusan Pendidikan Teknik Informatika," *KARMAPATI (Kumpulan Artikel Mahasiswa Pendidikan Teknik Informatika)*, vol. 6, no. 1, pp. 124–133, 2017.

[14] B. Suprapty and F. Fariyanti, "Klasifikasi Minat Siswa Untuk Program Studi Jurusan Teknologi Informasi-Politeknik Negeri Samarinda Menggunakan Metode Fuzzy C-means Clustering," *J-Icon: Jurnal Komputer dan Informatika*, vol. 8, no. 1, pp. 53–62, 2020.

[15] Y. S. Nugroho and A. K. Paleologoudias, "Differences Between Students from Senior High School and Vocational School in the Learning Outcomes of Electrical Engineering Students," in *2020 Third International Conference on Vocational Education and Electrical Engineering (ICVEE)*, IEEE, Oct. 2020, pp. 1–5. doi: 10.1109/ICVEE50212.2020.9243189.

[16] R. A. Cobb, "Vocational and Educational Aspirations of High School Students: A Problem for Rural America.," *Research in Rural Education*, vol. 6, no. 2, pp. 11–16, 1989.