Proceeding – CAMIC

Cirebon Annual Multidisciplinary International Conference (CAMIC 2024)

Optimization of Higher Education Management Policies through Big Data-Based Sentiment Analysis

1st 1st Asro
Master of Management
Unversitas Prmagraha
Banten, Indonesia
6110122023@upg.ac.id

2nd John Chaidir
Master of Management
Unversitas Prmagraha
Banten, Indonesia
johnchaidir@primagraha.ac.id

3rd Asep Ivan Permana
Master of Management
Unversitas Prmagraha
Banten, Indonesia
aseppermata2013@gmail.com

4th Khaled Fallaha
BSc IT Enterprise Computing
Skyline University
Sharjah, United Arab Emirates
khalidmfallaha@gmail.com

***Abstract*— This research utilizes big data and sentiment analysis to refine and optimize management policies in higher education sectors. It involves a comprehensive analysis of 52,181 comments gathered from social media across 16 provinces in Indonesia, focusing on stakeholders' perceptions regarding current educational policies. This study employs advanced analytical methods, including Naive Bayes, Support Vector Machine (SVM), and Logistic Regression, to process sentiments extracted from diverse digital interactions. The findings uncover a significant predominance of negative sentiment, accounting for 54% of the total comments. This indicates widespread dissatisfaction with the current educational frameworks. Positive sentiments account for 32.4%, and neutral sentiments make up 13.6%, suggesting areas where educational policies are received favorably, alongside points of contention. Among the analytical models employed, Logistic Regression achieved the highest accuracy at 99.81%, followed by SVM at 99.79% and Naive Bayes at 94.36%. This performance underscores the capability of data analysis technologies not only to manage but effectively interpret vast volumes of unstructured data, thus providing actionable insights. The results from this study highlight the critical need for adapting educational policies that align more closely with stakeholder expectations and regional educational dynamics, thereby improving the quality and reception of educational governance.**

Keywords— ***Sentiment Analysis; Big Data; Higher Education; Policy Optimization; Educational Management***

I. INTRODUCTION

In the current digital era, the utilization of big data and analytic technologies has revolutionized various aspects of industry, including higher education. Universities worldwide are pressured to adapt to rapid technological changes and enhance the quality of education and management policies. In this context, sentiment analysis emerges as a critical tool that enables institutions to understand stakeholders' perceptions and responses in real-time, fundamentally informing and reinforcing decision-making processes [1].

Sentiment analysis, a branch of data science that extracts opinions, attitudes, and emotions from text data, has been applied in various studies to describe and predict the complex dynamics of human interactions concerning diverse issues [2].

By leveraging data from social media and other interactive platforms, universities can gain valuable insights into the sentiments of students, faculty, and the public regarding institutional policies and practices [3],[4].

This approach is particularly useful for addressing the needs of students who may have insufficient perception and struggle with reading and comprehension strategies. By understanding these sentiments, educational institutions can tailor their interventions to enhance reading comprehension skills and deepen students' understanding of professionalism itself [5].

However, the effectiveness of sentiment analysis heavily depends on the algorithms used to process and analyze data [6]. Commonly employed algorithms include Naive Bayes, a classification method based on Bayes' theorem with an assumption of independence among predictors. This algorithm is renowned for its efficiency in processing large datasets with relatively quick computational times, making it ideal for real-time applications. Another frequently used algorithm is Support Vector Machines (SVM), which is effective in handling high-dimensional and complex datasets, such as text data from online comments or reviews [7].

Additionally, Logistic Regression is often utilized due to its capability in classifying binary data, which is extremely useful in sentiment analysis to assess the polarity (positive or negative) of texts. This study will explore both algorithms to determine the most effective in the context of sentiment analysis for higher education management policies[8].

With this background, the research aims to apply big data-based sentiment analysis techniques to optimize management policies at higher education institutions, with the hope of providing recommendations based on substantial and objective data analysis[9].

Implementing sentiment analysis for optimizing higher education management policies, the use of MongoDB as a NoSQL database system plays a crucial role in managing and analyzing big data [10]. MongoDB offers high flexibility with its document data model, which facilitates the storage of unstructured data such as texts from social media and discussion forums, primary sources for sentiment analysis[11].

Its dynamic schema feature allows universities to store data in various formats without the need to define a data structure upfront, highly beneficial given the volume and variability of data generated from digital platforms. Furthermore, MongoDB supports large data storage and high-performance real-time operations, essential for sentiment analysis that requires quick and efficient data processing. The use of big data in higher education not only enhances data analysis capabilities but also opens new opportunities in formulating policies that are more adaptive and responsive to the needs and expectations of stakeholders. Through the extensive data analysis gathered, universities can identify emerging sentiment patterns and respond with appropriate strategies that can improve student satisfaction and overall educational outcomes[8].

II. METHOD

This section outlines the research methodology utilized to conduct the study on optimizing higher education management policies through sentiment analysis. The methodological approach is depicted in the accompanying flowchart, which illustrates the sequential steps from data collection to decision-making. **Figure 1** below illustrates this methodology in detail.

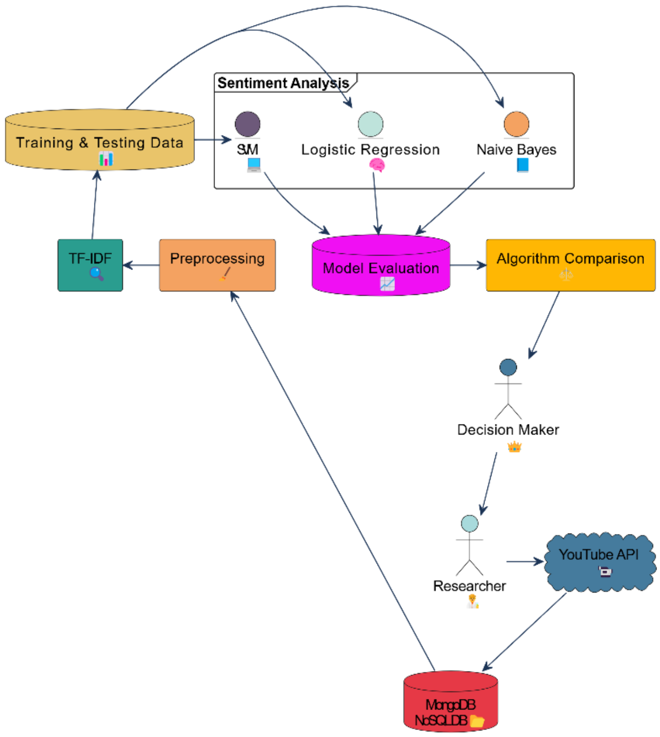

Figure I. **research methodology flowchart**

Description and Analysis:

The methodology employed in this study is designed to harness the power of big data and sentiment analysis for optimizing management policies in higher education. Each step in the process is crucial for ensuring the accuracy and relevancy of the findings.

1. Data Collection

The initial stage involves gathering a large volume of data directly from social media platforms to ensure a rich and diverse dataset. The **YouTube Data API** is employed to extract user comments, metadata, and timestamps relevant to higher education policy discussions. Key features of the dataset:

Source: YouTube comments related to higher education. Volume: 52,181 user comments. Coverage: Collected from 16 provinces in Indonesia. Storage: The data is stored in MongoDB, a NoSQL database system suitable for unstructured data, offering flexibility and scalability [12],[13].

2. Data Preprocessing

Preprocessing is a critical step to clean and prepare the dataset for analysis. This stage ensures that the subsequent analysis is based on high-quality and relevant data.

Preprocessing Steps:

a) Text Cleaning: Removal of special characters, emojis, and redundant spaces to standardize the text.
b) Normalization: Conversion of text to lowercase and application of stemming or lemmatization to unify word forms.
c) Stopword Removal: Filtering out common words that do not contribute to sentiment.
d) Vectorization: Transformation of text data into numerical format using TF-IDF (Term Frequency-Inverse Document Frequency)

3. Sentiment Analysis

The pre-processed data is subjected to sentiment analysis using three machine learning algorithms:
a) Naive Bayes: A probabilistic model based on Bayes' theorem, suitable for large datasets due to its simplicity and efficiency.
b) Support Vector Machine (SVM): A robust algorithm for handling high-dimensional data, offering high accuracy in sentiment classification.
c) Logistic Regression: A statistical model widely used for binary and multiclass classification tasks, providing excellent performance in sentiment polarity detection [14],[15].

4. Model Evaluation

The models are evaluated using predefined metrics to ensure their effectiveness in sentiment classification:
- Accuracy: Measures the proportion of correctly classified sentiments.
- Precision: Evaluates the proportion of true positives among all predicted positives.
- Recall: Measures the model's ability to identify relevant instances.
- F1-Score: Provides a balanced metric by combining precision and recall.

These metrics offer a comprehensive and systematic view of each model's performance, allowing for a fair and accurate comparison. The evaluation ensures that the selected model is well-suited for the specific requirements of sentiment classification in the context of higher education policy analysis

5. Algorithm Comparison

The comparison of algorithms provides a comprehensive understanding of the strengths and weaknesses of each method in sentiment analysis. This step is vital to selecting the most suitable algorithm for the task:

- Naive Bayes: Demonstrates simplicity and efficiency, achieving an accuracy of 94.36%. However, it struggles with distinguishing neutral sentiments, indicating a limitation in handling more nuanced data.

- Support Vector Machine (SVM): Achieves an accuracy of 99.79%, excelling in precision and recall across all sentiment classes. It is particularly effective for high-dimensional datasets, making it suitable for complex sentiment classification tasks.

- Logistic Regression: Attains the highest accuracy of 99.81%. It provides balanced performance across all metrics, making it the most reliable and effective model for this study.

This comparative analysis highlights the robustness of SVM and Logistic Regression, both of which outperform Naive Bayes in handling the complexities of sentiment analysis.

6. Decision Making

The final stage utilizes insights from sentiment analysis to guide and optimize higher education policy management. This step connects analytical results with actionable recommendations, closing the research loop.
- Addressing Dissatisfaction: The predominance of negative sentiments (54%) highlights dissatisfaction, requiring immediate policy interventions to address public concerns.
- Leveraging Positive Feedback: The 32.4% positive sentiments reflect areas where policies are effective and can be expanded or adapted to other regions.
- Targeted Regional Strategies: Sentiment distribution by province indicates that regions such as Sumatera Utara require focused efforts due to higher engagement and significant public discourse.

III. RESULTS AND DISCUSSION

In their 2024 study, Henderi et al. utilized sentiment analysis to assess public perceptions of education in Indonesia by analyzing YouTube comments. They processed 13,386 comments, revealing that 53.8% were negative, 28.5% positive, and 17.7% neutral, indicating diverse viewpoints on educational issues [16]. Similarly, Bawa et al. (2024) examined public sentiment toward the "Merdeka Belajar Kampus Merdeka" (MBKM) policy using Twitter data. Employing the K-Nearest Neighbor method, they analyzed 2,000 tweets and found that negative sentiments comprised 85%, neutral 53%, and positive 70%, with an overall accuracy of 71%. These findings suggest a predominantly negative public perception of the MBKM policy.

In their 2023 study, Padri et al. utilized the Naive Bayes method to analyze Twitter data for detecting traffic congestion in Indonesia. They found that this approach effectively identified relevant information about traffic jams, aiding in strategic planning to address congestion issues [17].

In their 2024 study, Asro et al. employed MongoDB to analyze 27,102 YouTube comments related to the Indonesian Gubernatorial Election. They found that Logistic Regression achieved a 91.39% accuracy in sentiment analysis, highlighting MongoDB's effectiveness in processing large-scale social media data[18].

In their study, Asro et al. (2025) evaluated the performance of Decision Tree, K-Nearest Neighbors (KNN), and Naive Bayes algorithms in predicting student graduation at Universitas XYZ. The Decision Tree algorithm achieved the highest accuracy at 97%, followed closely by KNN, while Naive Bayes recorded an accuracy of 91%. The researchers highlighted the Decision Tree's ease of interpretation and transparency, making it particularly suitable for educational applications. These findings offer valuable insights for educational institutions aiming to develop effective graduation prediction systems and inform strategic academic management decisions[19].

Building upon these methodologies, this study collected 52,181 comments from social media platforms across 16 Indonesian provinces, focusing on discussions related to higher education policies. The analysis aims to provide insights into public sentiment and inform policy-making in the educational sector.

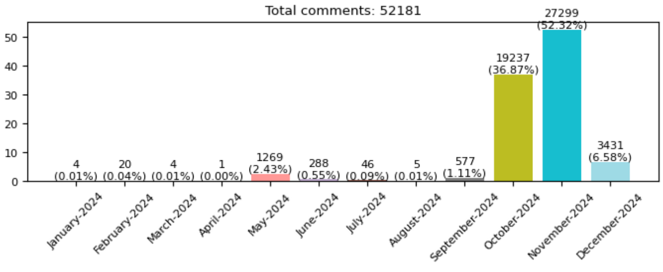

Figure 2. **Data Collection Monthly**

Data Collection Overview: The data collection spanned from January to December 2024, capturing a broad temporal snapshot of public sentiment towards higher education policies. The graph below provides a monthly distribution of the comments, highlighting significant variances in the volume of discourse throughout the year.

January to April: During these months, the discussion volume was notably low, with totals ranging from just 4 comments in January to 46 in July. This period may reflect a lack of significant policy changes or public events related to higher education that would typically spur increased social media activity.

May: A slight increase occurred in May with **1,269** comments, possibly due to specific policy announcements or developments in the educational sector that captured public interest.

June to September: The volume remained relatively low during these months, with September only seeing **577** comments, suggesting that there may not have been major contentious issues or changes during this period.

October to December: A substantial spike in discussion was observed starting in October, with 19,237 comments, and peaking in November with **27,299** comments. This surge could be attributed to heated discussions related to year-end educational assessments, budget allocations, or policy revisions that typically occur during this time.

Analysis of Comment Distribution: The bulk of the comments in the latter part of the year, particularly in November and December, indicates a heightened level of engagement from the public concerning higher education policies. The intense discussions during these months likely reflect critical evaluations of the year's policies, debates over proposed changes, and reactions to new policy implementations announced towards the year's end.

Distribution of Comments by Province: The geographical distribution of the comments collected from social media platforms is represented in Figure 3. The pie chart illustrates the proportion of comments contributed by each province, highlighting variations in engagement levels across Indonesia.

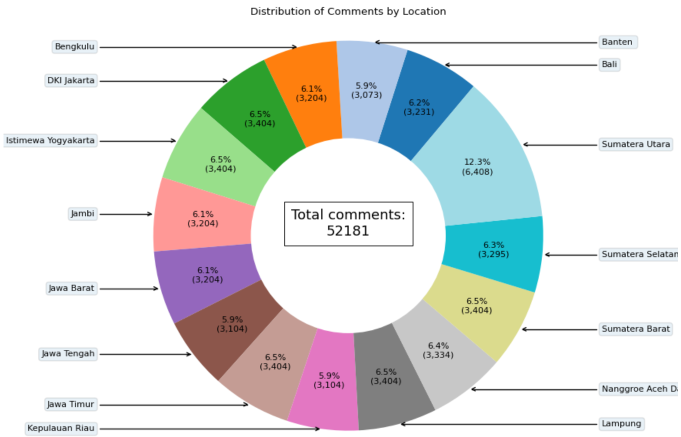

Figure 3. **Distribution of Comments by Province**

Overview of the Pie Chart:

Sumatera Utara: This province shows the highest engagement, contributing 12.3% of the total comments (6,408 comments). This significant percentage indicates a robust interaction and concern about educational policies within the region.

Other Notable Contributions: DKI Jakarta and Jawa Timur: Both regions also show substantial engagement, each contributing 6.5% of the comments. Given the population density and educational infrastructure in these areas, the high level of engagement is expected.

Sumatera Selatan and Sumatera Barat: Similar to DKI Jakarta and Jawa Timur, these provinces each contributed 6.4% and 6.3% respectively, reflecting active discussions around educational policies.

Regions with Lower Engagement: Bali and Banten: These areas showed relatively lower engagement, contributing 5.9% of the comments each. Despite being smaller regions, their active participation underscores a widespread interest in educational policy across diverse geographic areas.

Interpretation and Implications: This data distribution is pivotal for discerning regional variances in how higher education policies are perceived and reacted to. The heightened engagement from regions such as Sumatera Utara highlights zones where educational policies might be more contentious or impactful. Conversely, a consistent distribution across numerous provinces reflects a broad national interest in the discourse and reform of educational policies. These insights are supported by technologies that have been demonstrated to foster active learning and enhance student satisfaction [20].

Sentiment Analysis Overview After mapping out the geographical distribution, a detailed sentiment analysis was conducted. The sentiments were categorized into negative, positive, and neutral. As summarized in Figure 4, the analysis reveals a predominant negative sentiment, accounting for 54.0% of the comments, suggesting widespread dissatisfaction with the current educational policies. Positive and neutral sentiments accounted for 32.4% and 13.6%, respectively, indicating some areas of approval and ambivalence.

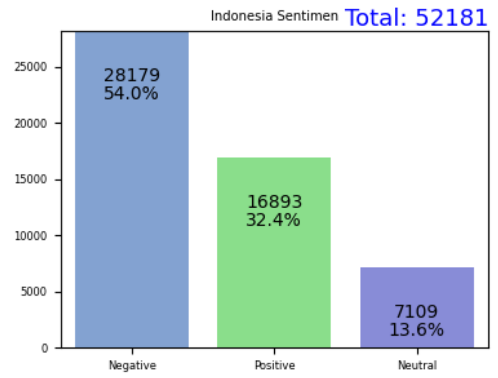

Figure 4. Sentiment Distribution

Model Evaluation:

The effectiveness of the sentiment analysis models used in the study was evaluated by applying three distinct algorithms: Naive Bayes, Support Vector Machine (SVM), and Logistic Regression. Each model's performance was assessed based on its ability to accurately classify sentiments, which is essential for understanding and addressing public opinions on higher education policies.

Naive Bayes Model Performance Overview Accuracy: 94.36% The Naive Bayes classifier has demonstrated a commendable performance with an accuracy of 94.36%. This model, popular for its simplicity in handling text classification tasks, shows substantial proficiency, especially in identifying sentiments from text data. Confusion Matrix Visualization and Analysis[21].

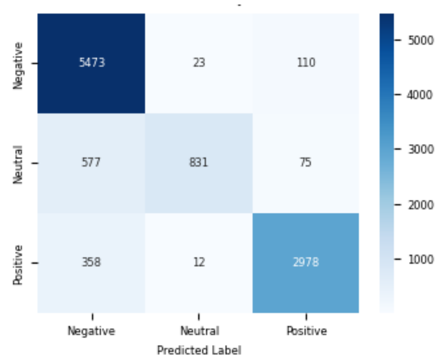

Figure 5. Confusion Matrix for Naive Bayes

The confusion matrix above offers a detailed view of the model's performance across different sentiment classes: Negative, Neutral, and Positive. Here's how the model performed for each class:

- Negative Sentiments: High recall (0.98) implies that the model is very effective in identifying negative comments. However, the precision of 0.85 suggests some instances of other sentiments being misclassified as negative.
- Neutral Sentiments: The low recall (0.56) indicates a challenge in accurately identifying neutral comments, which are often confused with other sentiments.
- Positive Sentiments: High values in both precision (0.94) and recall (0.89) indicate strong performance in recognizing positive sentiments.

Manual Calculation of Naïve Bayes Metrics:

a) $Accuracy = \frac{TP + TN}{Total} = \frac{2978 + 6304}{9837} x100\% = 94.36\%$

b) $Precision = \frac{TP}{TP + FP} = \frac{2978}{2978 + 185} = 0.942$

c) $Recal = \frac{TP}{TP + FN} = \frac{2978}{2978 + 370} = 0.889$

d) $F1\,SCORE = 2\,X \left(\frac{Precision\ X\ Recall}{Precision + Recall}\right) = 2\,X \left(\frac{0.942\ X\ 0.889}{0.942 + 0.889}\right) = 0.915$

These metrics offer a comprehensive overview of the effectiveness of the Naive Bayes model in classifying sentiments, showing strengths in identifying positive and negative sentiments while noting that there is room for improvement in recognizing neutral sentiments.

SVM Model Performance Overview Accuracy: 99.79% The SVM classifier, known for its robustness in handling high-dimensional data, demonstrated exceptional accuracy in this study. Its capability to effectively separate positive from negative and neutral sentiments is particularly notable, making it invaluable for nuanced sentiment analysis. Confusion Matrix Visualization and Analysis [1].

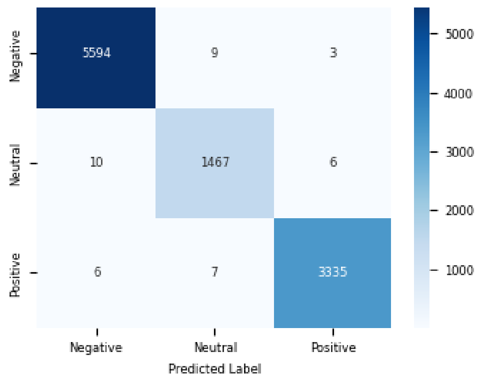

Figure 6. **Confusion Matrix for SVM**

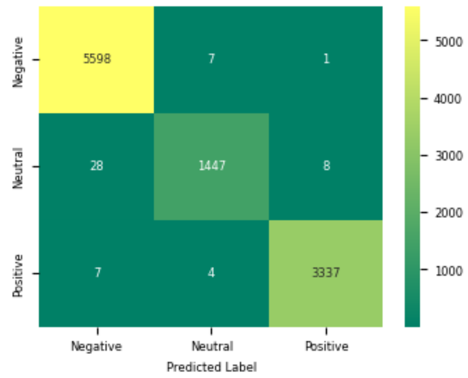

Figure 7. **Confusion Matrix for Logistic Regression**

The confusion matrix above provides a detailed view of the model's performance across different sentiment classes: Negative, Neutral, and Positive. Here's how the model performed for each class:

- Negative Sentiments: Perfect scores in precision and recall (1.00) indicate flawless classification of negative comments.
- Neutral Sentiments: Slightly lower, yet near-perfect recall (0.99) suggests that there are few misclassifications of neutral sentiments, likely as positive or negative.
- Positive Sentiments: Again, perfect scores in both precision and recall (1.00) indicate excellent recognition of positive sentiments.

Manual Calculation of SVM Metrics:

e) $Accuracy = \frac{TP+TN}{Total} = \frac{3335+7061}{10418} x100\% = 99.79\%$

f) $Precision = \frac{TP}{TP+FP} = \frac{3335}{3335+9} = 0.997$

g) $Recal = \frac{TP}{TP+FN} = \frac{3335}{3335+13} = 0.996$

h) $F1SCORE = 2\,X \left(\frac{Precision\ X\ Recall}{Precision + Recall}\right) = 2\,X \left(\frac{0.997\ X\ 0.996}{0.997+0.996}\right) = 0.997$

These metrics underscore the effectiveness of the SVM model in sentiment classification, demonstrating its strengths in accurately categorizing negative, neutral, and positive sentiments with remarkable precision and recall, specifically highlighting its prowess in nuanced detection tasks.

Logistic Regression Model Performance Overview Accuracy: 99.81% Logistic Regression has shown the highest accuracy among the models tested, excelling in distinguishing between all sentiment categories with minimal error rates. Its robust performance highlights its suitability for complex classification tasks where precision and reliability are paramount. Confusion Matrix Visualization and Analysis.

This confusion matrix provides a clear view of the Logistic Regression model's performance across different sentiment classes. Here's how the model performed for each class:

- High Accuracy: Almost perfect scores across the board for all classes demonstrate the model's excellent ability to correctly classify sentiments.
- Precision and Recall: Nearly identical high scores for both metrics indicate a very balanced and accurate model.

Manual Calculation of logistic regression Metrics:

i) $Accuracy = \frac{TP+TN}{Total} = \frac{3337+7045}{10402} x100\% = 99.81\%$

j) $Precision = \frac{TP}{TP+FP} = \frac{3337}{3337+9} = 0.997$

k) $Recal = \frac{TP}{TP+FN} = \frac{3337}{3337+11} = 0.997$

l) $F1SCORE = 2\,X \left(\frac{Precision\ X\ Recall}{Precision + Recall}\right) = 2\,X \left(\frac{0.997\ X\ 0.997}{0.997+0.997}\right) = 0.997$

Logistic Regression is a statistical model that examines the relationship between independent variables—here, words in text—and a binary dependent variable. It classifies data into categories; in this context, three sentiment classes: Negative, Neutral, and Positive. The multinomial Logistic Regression function is expressed as follows:

a) $gj(x) = \beta j0 + \beta j1x1 + \beta j2x2 + \cdots + \beta jpxp$

b) *Description:*

c) $g(x) = Logit\ function\ for\ class\ j$

d) $\beta j0 = intercept\ for\ class\ \boldsymbol{j}$

e) $\beta j1, \beta j2, \ldots, \beta jp$ =*coefficient for each predictor variable* $(x1, x2, \ldots, xp)$ *in class j*

Figure 8: Screenshot Logistic Regression Formula

These metrics confirm the exceptional performance of the Logistic Regression model in sentiment classification, demonstrating its strengths in accurately categorizing negative, neutral, and positive sentiments with almost perfect precision and recall. This makes it particularly effective for nuanced detection tasks in varied data environments.

Discussion and Policy Implications:

The sentiment analysis reveals a predominant negative sentiment across the dataset, indicating a general dissatisfaction with existing higher education policies. The models' high accuracy rates, especially in the SVM and Logistic Regression, lend confidence to these findings and suggest specific areas for policy intervention.

The regional variations in sentiment highlighted in the sentiment distribution and the detailed model evaluations suggest that local factors significantly influence public opinion. These insights can guide policymakers to tailor educational reforms more effectively to regional needs and conditions, potentially enhancing public satisfaction and policy effectiveness.

The models' comparative analysis highlights the strengths of machine learning approaches in processing and analyzing vast amounts of unstructured data. Policymakers are encouraged to leverage these technologies to continuously monitor and adapt to public sentiment, ensuring that educational policies remain aligned with the stakeholders' evolving needs.

IV. CONCLUSIONS

This research has effectively demonstrated the application of big data and sentiment analysis in optimizing higher education management policies. Through the detailed examination of 52,181 comments extracted from social media platforms across 16 provinces in Indonesia, the study provides a data-driven insight into public sentiments regarding current educational policies. Key Findings:

- Predominant Negative Sentiment: The analysis revealed that a significant 54.0% of the sentiments expressed were negative, indicating widespread dissatisfaction with the existing educational policies.
- Regional Variations: The distribution of comments highlighted that certain region, such as Sumatera Utara, were particularly vocal in their feedback, suggesting a regional specificity to the concerns and opinions about higher education.
- Model Effectiveness: Among the models utilized, Logistic Regression and SVM demonstrated superior performance with accuracies of 99.81% and 99.79% respectively, proving their efficacy in accurately classifying complex sentiment data.
- The application of advanced analytical models has not only facilitated a deeper understanding of the public's opinion at a granular level but also highlighted the potential areas for policy intervention. By leveraging the insights from this sentiment analysis, policymakers can prioritize areas that require urgent attention and customize responses to regional educational needs, thereby enhancing the effectiveness and acceptance of education policies.

REFERENCES

[1] K. Adib, M. R. Handayani, W. D. Yuniarti, And K. Umam, “Opini Publik Pasca-Pemilihan Presiden: Eksplorasi Analisis Sentimen Media Sosial X Menggunakan Svm,” Aug. 2024, [Online]. Available: Https://Doi.Org/10.31598

[2] P. Nandwani And R. Verma, “A Review On Sentiment Analysis And Emotion Detection From Text,” Dec. 01, 2021, *Springer*. Doi: 10.1007/S13278-021-00776-6.

[3] K. Wimshurst, R. Wortley, M. Bates, And T. Allard, “The Impact Of Institutional Factors On Student Academic Results: Implications For ‘Quality’ In Universities,” *Higher Education Research And Development*, Vol. 25, No. 2, Pp. 131–145, 2006, Doi: 10.1080/07294360600610370.

[4] J. Friadi *Et Al.*, “E-Learning Evaluation Based On Context, Input, Process, And Product (Cipp),” *Jurnal Penelitian Pendidikan Ipa*, Vol. 10, No. Specialissue, Pp. 67–73, Aug. 2024, Doi: 10.29303/Jppipa.V10ispecialissue.7383.

[5] B. Wibisono, “Perception Of Doctor Professionalism In Medical Students.”

[6] A. Robi Padri, A. Asro, And I. Indra, “Classification Of Traffic Congestion In Indonesia Using The Naive Bayes Classification Method,” *Journal Of World Science*, Vol. 2, No. 6, Pp. 877–888, Jun. 2023, Doi: 10.58344/Jws.V2i6.285.

[7] S. W. Iriananda, R. W. Budiawan, A. Y. Rahman, And I. Istiadi, “Optimasi Klasifikasi Sentimen Komentar Pengguna Game Bergerak Menggunakan Svm, Grid Search Dan Kombinasi N-Gram,” *Jurnal Teknologi Informasi Dan Ilmu Komputer*, Vol. 11, No. 4, Pp. 743–752, Aug. 2024, Doi: 10.25126/Jtiik.1148244.

[8] R. Baragash And H. Aldowah, “Sentiment Analysis In Higher Education: A Systematic Mapping Review,” In *Journal Of Physics: Conference Series*, Iop Publishing Ltd, Apr. 2021. Doi: 10.1088/1742-6596/1860/1/012002.

[9] J. Chaidir And H. Haerofiatna, “Development Of Data Management In The Implementation Of Electronic-Based Government In Serang Regency,” *International Journal Of Community And Cooperative Studies*, Vol. 11, No. 1, Pp. 17–25, Jan. 2023, Doi: 10.37745/Ijccs.2014/Vol11n11725.

[10] S. Kumari Behera *Et Al.*, “Mongodb Integration With Python And Node.Js, Express.Js,” *2024 Fourth International Conference On Advances In Electrical, Computing, Communication And Sustainable Technologies*, 2024, Doi: 10.1109/Icaect60202.2024.10460546.

[11] R. K. Mohan, R. R. S. Kanmani, K. A. Ganesan, And N. Ramasubramanian, “Evaluating Nosql Databases For Olap Workloads: A Benchmarking Study Of Mongodb, Redis, Kudu And Arangodb,” May 2024, [Online]. Available: Http://Arxiv.Org/Abs/2405.17731

[12] A. Neelkant, A. Rawat, A. Anilkumar, A. Sarmah, And A. Jadhav, “Sentimental Analysis On Twitter,” 2024.

[13] R. S. Al Fathir As, E. Utami, And A. Dwi Hartono, “Comparison Of Sentiment Analysis Methods On Topic Haram Of Music In Youtube,” *Sinkron*, Vol. 8, No. 4, Pp. 2294–2307, Oct. 2023, Doi: 10.33395/Sinkron.V8i4.12776

[14] K. Adib, M. R. Handayani, W. D. Yuniarti, And K. Umam, “Opini Publik Pasca-Pemilihan Presiden:

Eksplorasi Analisis Sentimen Media Sosial X Menggunakan Svm," Aug. 2024, [Online]. Available: Https://Doi.Org/10.31598

[15] P. Nandwani And R. Verma, "A Review On Sentiment Analysis And Emotion Detection From Text," Dec. 01, 2021, *Springer*. Doi: 10.1007/S13278-021-00776-6.

[16] K. Wimshurst, R. Wortley, M. Bates, And T. Allard, "The Impact Of Institutional Factors On Student Academic Results: Implications For 'Quality' In Universities," *Higher Education Research And Development*, Vol. 25, No. 2, Pp. 131–145, 2006, Doi: 10.1080/07294360600610370.

[17] J. Friadi *Et Al.*, "E-Learning Evaluation Based On Context, Input, Process, And Product (Cipp)," *Jurnal Penelitian Pendidikan Ipa*, Vol. 10, No. Specialissue, Pp. 67–73, Aug. 2024, Doi: 10.29303/Jppipa.V10ispecialissue.7383.

[18] B. Wibisono, "Perception Of Doctor Professionalism In Medical Students."

[19] A. Robi Padri, A. Asro, And I. Indra, "Classification Of Traffic Congestion In Indonesia Using The Naive Bayes Classification Method," *Journal Of World Science*, Vol. 2, No. 6, Pp. 877–888, Jun. 2023, Doi: 10.58344/Jws.V2i6.285.

[20] S. W. Iriananda, R. W. Budiawan, A. Y. Rahman, And I. Istiadi, "Optimasi Klasifikasi Sentimen Komentar Pengguna Game Bergerak Menggunakan Svm, Grid Search Dan Kombinasi N-Gram," *Jurnal Teknologi Informasi Dan Ilmu Komputer*, Vol. 11, No. 4, Pp. 743–752, Aug. 2024, Doi: 10.25126/Jtiik.1148244.

[21] R. Baragash And H. Aldowah, "Sentiment Analysis In Higher Education: A Systematic Mapping Review," In *Journal Of Physics: Conference Series*, Iop Publishing Ltd, Apr. 2021. Doi: 10.1088/1742-6596/1860/1/012002.

[22] J. Chaidir And H. Haerofiatna, "Development Of Data Management In The Implementation Of Electronic-Based Government In Serang Regency," *International Journal Of Community And Cooperative Studies*, Vol. 11, No. 1, Pp. 17–25, Jan. 2023, Doi: 10.37745/Ijccs.2014/Vol11n11725.

[23] S. Kumari Behera *Et Al.*, "Mongodb Integration With Python And Node.Js, Express.Js," *2024 Fourth International Conference On Advances In Electrical, Computing, Communication And Sustainable Technologies*, 2024, Doi: 10.1109/Icaect60202.2024.10460546.

[24] R. K. Mohan, R. R. S. Kanmani, K. A. Ganesan, And N. Ramasubramanian, "Evaluating Nosql Databases For Olap Workloads: A Benchmarking Study Of Mongodb, Redis, Kudu And Arangodb," May 2024, [Online]. Available: Http://Arxiv.Org/Abs/2405.17731

[25] A. Neelkant, A. Rawat, A. Anilkumar, A. Sarmah, And A. Jadhav, "Sentimental Analysis On Twitter," 2024.

[26] R. S. Al Fathir As, E. Utami, And A. Dwi Hartono, "Comparison Of Sentiment Analysis Methods On Topic Haram Of Music In Youtube," *Sinkron*, Vol. 8, No. 4, Pp. 2294–2307, Oct. 2023, Doi: 10.33395/Sinkron.V8i4.12776.

[27] A. Sulaiman, "Analisis Sentimen Tentang Transformasi Program Makan Siang Menjadi Makan Bergizi Gratis Menggunakan Logistik Regression Pada Laman Youtube," *Jurnal Ict : Information Communication & Technology*, Vol. 24, No. 1, 2024, [Online]. Available: Https://Ejournal.Ikmi.Ac.Id/Index.Php/Jict-Ikmi

[28] A. N. S. Asro, "Public Sentiment Analysis On The 2024 Presidential Election Using Naive Bayes Classifier (Nbc) And Support Vector Machine (Svm) On Social Media Data," *Prosiding Seminar Internasional*, 2024, [Online]. Available: Https://Prosiding.Amalinsani.Org/

[29] Henderi, Asro, A. Sulaiman, T. B. Kurniawan, D. A. Dewi, And M. Alqudah, "Utilizing Sentiment Analysis For Reflect And Improve Education In Indonesia," *Journal Of Applied Data Sciences*, Vol. 6, No. 1, Pp. 189–200, Jan. 2025, Doi: 10.47738/Jads.V6i1.527.

[30] A. Robi Padri, A. Asro, And I. Indra, "Classification Of Traffic Congestion In Indonesia Using The Naive Bayes Classification Method," *Journal Of World Science*, Vol. 2, No. 6, Pp. 877–888, Jun. 2023, Doi: 10.58344/Jws.V2i6.285.

[31] A. Hartono Limaran, A. Wicaksono, And P. Herwanto, "Strategi Dinamis Menggunakan Mongodb Untuk Analisis Sentimen Terhadap Komentar Youtube Pilkada Gubernur Indonesia 2024 Dynamic Strategy Using Mongodb For Sentiment Analysis Of Youtube Comments On The 2024 Indonesian Gubernatorial Elections." [Online]. Available: Https://Jurnaldrpm.Budiluhur.Ac.Id/Index.Php/Kresna/

[32] J. Chaidir And J. Friadi, "Zona Teknik: Jurnal Ilmiah Evaluasi Kinerja Algoritma Klasifikasi Dalam Studi Kasus Prediksi Kelulusan Di Universitas Xyz," Vol. 19, No. 1, Pp. 15–22, 2025, Doi: 10.37776/Zt.Vxix.Xxx.

[33] P. Fasco, S. Asiimwe, T. Atwongire, P. Shatson Fasco, G. Ssekabira, And J. A. Tushabe, "Enhancing Student Engagement And Learning Outcomes: Effective Strategies In Institutional Pedagogy", [Online]. Available: Www.Allmultidisciplinaryjournal.Com

[34] D. Nurcahyono, W. P. Putra, A. Najib, And T. R. Tulili, "Analysis Sentiment In Social Media Against Election Using The Method Naive Bayes," In *Journal Of Physics: Conference Series*, Institute Of Physics Publishing, Jun. 2020. Doi: 10.1088/1742-6596/1511/1/012003.