Journal of Applied Data Sciences
Vol. 6, No. 2, May 2025, pp. 1076-1087

Sentiment Analysis on Slang Enriched Texts Using Machine Learning Approaches

Priyo Agung Prastyo[1,*], Berlilana[2], Imam Tahyudin[3]

[1,2,3] *Magister of Computer Science, Universitas Amikom Purwokerto, Jawa Tengah, Indonesia*

(Received: November 13, 2024; Revised: December 19, 2024; Accepted: January 17, 2025; Available online: March 4, 2025)

Abstract

This study explores sentiment analysis of slang-enriched user reviews using machine learning techniques, specifically Naive Bayes, Support Vector Machine (SVM), and Random Forest, to classify user sentiment into Positive, Negative, and Neutral categories while addressing challenges posed by informal and conversational language through slang normalization. A lexicon-based scoring method was employed to standardize slang terms such as “gak,” “aja,” and “banget,” ensuring consistency in sentiment analysis. The results indicate that Neutral sentiment dominates the dataset (51%), followed by Negative (28%) and Positive (21%), with lexicon-based scores confirming this distribution. Negative sentiment exhibits a broader intensity range, reflecting user dissatisfaction primarily related to network quality, service reliability, and pricing, as evident from recurring terms like “sinyal” (signal), “jaringan” (network), and “mahal” (expensive). Word cloud visualizations reinforce these findings, highlighting the prevalence of these concerns in user feedback. Performance evaluation of the machine learning models reveals that SVM and Random Forest achieved the highest accuracy (96%), significantly outperforming Naive Bayes (73%), demonstrating their effectiveness in handling high-dimensional text data and accurately classifying slang-rich content. These findings underscore the importance of slang normalization in preprocessing, as it significantly enhances sentiment classification accuracy. This study provides actionable insights for service providers, helping them identify and address key sources of user dissatisfaction. Future research can explore deep learning models such as BERT and LSTM to further enhance sentiment analysis by capturing contextual relationships within text data, while topic modeling techniques could uncover deeper thematic patterns in user feedback, enabling data-driven strategies to improve customer satisfaction.

Keywords: Sentiment Analysis, Slang Normalization, Lexicon-Based Scoring, Machine Learning Models, Support Vector Machine (SVM), Random Forest, User Reviews Analysis

1. Introduction

Sentiment analysis has become a critical tool for understanding user opinions and extracting insights from textual data, particularly in the context of product reviews, customer feedback, and social media content [1], [2]. By analyzing text and classifying it into sentiment categories such as Positive, Negative, or Neutral, organizations can gauge user satisfaction, identify key areas for improvement, and make data-driven decisions [3], [4]. However, sentiment analysis becomes significantly more challenging when dealing with informal and slang-enriched language, which is increasingly prevalent in user-generated content [5]. This challenge arises from the unstructured nature of such texts, the frequent use of regional slang, abbreviations, and colloquial expressions, which often distort the sentiment meaning and hinder model performance [6].

Slang language, characterized by its informal tone and non-standard vocabulary, poses unique challenges for traditional natural language processing (NLP) techniques [7]. For instance, words like *“gak”* (meaning “no”) and *“banget”* (meaning “very”) are widely used in Indonesian conversational contexts but are not part of formal language corpora [8]. Without proper preprocessing and normalization, these slang terms may be misinterpreted or excluded during text processing, leading to inaccurate sentiment classification [9]. Therefore, slang normalization becomes a crucial step in preparing such data for sentiment analysis. This study incorporates a predefined slang dictionary to standardize informal terms, ensuring that slang does not compromise the sentiment analysis process [10]. The integration of lexicon-based scoring and machine learning approaches offers a robust framework for addressing this challenge [11].

***Corresponding author: Priyo Agung Prastyo (priyoagung88@gmail.com)**
DOI: https://doi.org/10.47738/jads.v6i2.626
This is an open access article under the CC-BY license (https://creativecommons.org/licenses/by/4.0/).
© Authors retain all copyrights

Lexicon-based methods rely on predefined dictionaries where words are assigned sentiment scores based on their polarity, providing an interpretable approach to assigning sentiment to text [12]. While lexicon-based methods are useful for initial scoring, they may struggle with complex patterns in informal text [13]. To overcome this limitation, machine learning models such as Naive Bayes, Support Vector Machine (SVM), and Random Forest are employed to classify sentiment [14]. Naive Bayes serves as a probabilistic baseline model due to its computational efficiency, while SVM and Random Forest are capable of handling high-dimensional features and complex relationships within the data, making them suitable for analyzing slang-enriched user reviews [15].

In recent years, sentiment analysis research has focused on formal texts, such as structured reviews or news articles [16]. However, limited attention has been given to texts containing informal language and slang, which are prevalent in user-generated reviews [1]. This research addresses this gap by applying sentiment analysis to a dataset of user reviews written in conversational Indonesian, enriched with slang expressions [2]. The study focuses on three key objectives: (1) to preprocess slang-enriched text using slang normalization [3], (2) to classify sentiment using lexicon-based scoring and machine learning models [4], and (3) to evaluate the performance of Naive Bayes, Support Vector Machine (SVM), and Random Forest models in handling such data [5].

The significance of this research lies in its ability to extract actionable insights from user reviews that are often difficult to process due to the prevalence of informal language [6]. By addressing the challenges posed by slang and conversational expressions, this study provides a systematic methodology for accurately classifying sentiment and identifying recurring user concerns [7]. The findings have practical implications for service providers and businesses, as they highlight specific areas of dissatisfaction, such as network quality, service reliability, and pricing issues [8], while also emphasizing opportunities to improve user satisfaction [9].

In this study, we propose a comprehensive framework that combines slang normalization, lexicon-based scoring, and machine learning modeling to analyze user sentiments [10]. The study begins with text preprocessing, including the identification and normalization of slang words [11], followed by lexicon-based sentiment scoring and feature extraction using TF-IDF [12]. The processed text is then classified into sentiment categories using Naive Bayes, SVM, and Random Forest models, whose performance is evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score [13]. Additional analyses, such as word cloud visualization and slang frequency analysis, are conducted to provide deeper insights into the themes and topics discussed by users [14].

By combining robust preprocessing techniques with advanced machine learning methods, this study addresses the challenges of analyzing slang-enriched text and contributes to the growing field of sentiment analysis [15]. The proposed methodology not only enhances the accuracy of sentiment classification but also provides valuable insights into user feedback, enabling businesses and service providers to make informed decisions and improve customer satisfaction [16].

2. Literature Review

Sentiment analysis, a subfield of Natural Language Processing (NLP), has gained significant attention in recent years as a method for analyzing textual data to extract opinions and classify them into sentiment categories such as Positive, Negative, or Neutral [17]. The ability to derive insights from large volumes of user-generated content, such as reviews and social media posts, has proven valuable for businesses, policymakers, and researchers [18]. Traditional sentiment analysis techniques rely on rule-based or lexicon-based approaches, where predefined word dictionaries are used to assign sentiment scores to text [19]. These methods are interpretable and effective for structured, formal texts but often fail to handle unstructured and informal language, such as slang and abbreviations [20].

The use of slang and conversational language introduces challenges that require additional preprocessing steps. Slang words, which are often context-dependent, vary significantly across regions and languages [21]. For example, words like *“gak”* (meaning “no”) or *“banget”* (meaning “very”) in Indonesian are widely used but are not part of standard dictionaries. This discrepancy can distort the output of traditional sentiment analysis models. Several studies have highlighted the importance of slang normalization to replace informal words with their formal equivalents before further analysis [22]. The process typically involves the use of predefined slang dictionaries or context-aware models to ensure consistency in text representation. Studies have demonstrated that proper normalization significantly enhances the accuracy of both lexicon-based and machine learning-based sentiment classifiers [23].

To overcome the limitations of rule-based methods, machine learning approaches have been widely adopted for sentiment classification. Machine learning models, such as Naive Bayes, Support Vector Machine (SVM), and Random Forest, have demonstrated superior performance in handling unstructured textual data [24]. Naive Bayes, a probabilistic classifier, is computationally efficient and widely used for text classification tasks due to its simplicity. However, it assumes independence between features, which may limit its effectiveness when analyzing complex patterns in slang-enriched text. On the other hand, SVM and Random Forest are robust classifiers capable of handling high-dimensional data and nonlinear relationships. SVM identifies optimal hyperplanes for separating sentiment classes, while Random Forest aggregates predictions from multiple decision trees to reduce overfitting and improve accuracy [25]. These models are particularly effective in scenarios where the text contains significant noise, such as slang or informal expressions.

Feature extraction plays a crucial role in transforming textual data into a format suitable for machine learning. The TF-IDF (Term Frequency-Inverse Document Frequency) method is one of the most commonly used techniques for representing text as numerical features. TF-IDF assigns higher importance to words that appear frequently in a specific document but are rare across the overall dataset, enabling the model to focus on terms that contribute most to sentiment classification [17]. In the context of slang-enriched text, TF-IDF can capture normalized slang terms after preprocessing, ensuring that informal expressions are appropriately weighted. Recent advancements have also explored the use of word embeddings, such as Word2Vec and GloVe, to represent text in a continuous vector space and capture semantic relationships between words [18].

While traditional machine learning models perform well on sentiment analysis tasks, several studies have shown that the integration of deep learning techniques can further enhance performance. Models like Long Short-Term Memory (LSTM) networks and Bidirectional Encoder Representations from Transformers (BERT) have the ability to capture contextual relationships within the text, making them highly effective for analyzing slang-enriched and unstructured content [19], [20]. For example, LSTM networks can process sequences of words while maintaining temporal dependencies, while BERT leverages bidirectional context to understand the sentiment polarity of words in relation to surrounding terms. However, the computational complexity of deep learning models often necessitates large datasets and significant resources, which may not always be feasible in smaller-scale studies.

Previous research has primarily focused on sentiment analysis of formal texts, such as structured product reviews, news articles, and academic texts [21]. Limited studies have explored the challenges posed by informal, slang-heavy content, particularly in regional languages like Indonesian [22]. The growing use of slang in online reviews and social media highlights the need for tailored preprocessing techniques and robust machine learning frameworks capable of handling such content. This study builds on existing research by addressing the challenges of slang normalization, integrating lexicon-based scoring with machine learning classifiers, and evaluating their performance on a dataset of user-generated reviews.

In conclusion, while lexicon-based methods offer simplicity and interpretability, machine learning models such as Naive Bayes, SVM, and Random Forest provide superior performance when analyzing complex and noisy text data [23], [24]. The integration of preprocessing techniques, particularly slang normalization, remains critical for improving the accuracy of sentiment classification in informal texts. This study extends prior work by implementing a comprehensive framework that combines slang normalization, lexicon-based scoring, and machine learning-based classification, providing new insights into the challenges and opportunities of sentiment analysis in slang-enriched user reviews [25].

3. Methodology

The methodology of this study involves a structured and systematic approach to analyze user reviews containing slang-enriched text for sentiment classification as shown in figure 1. This approach integrates text preprocessing, sentiment scoring, and machine learning modeling to classify the reviews into Positive, Negative, and Neutral categories.

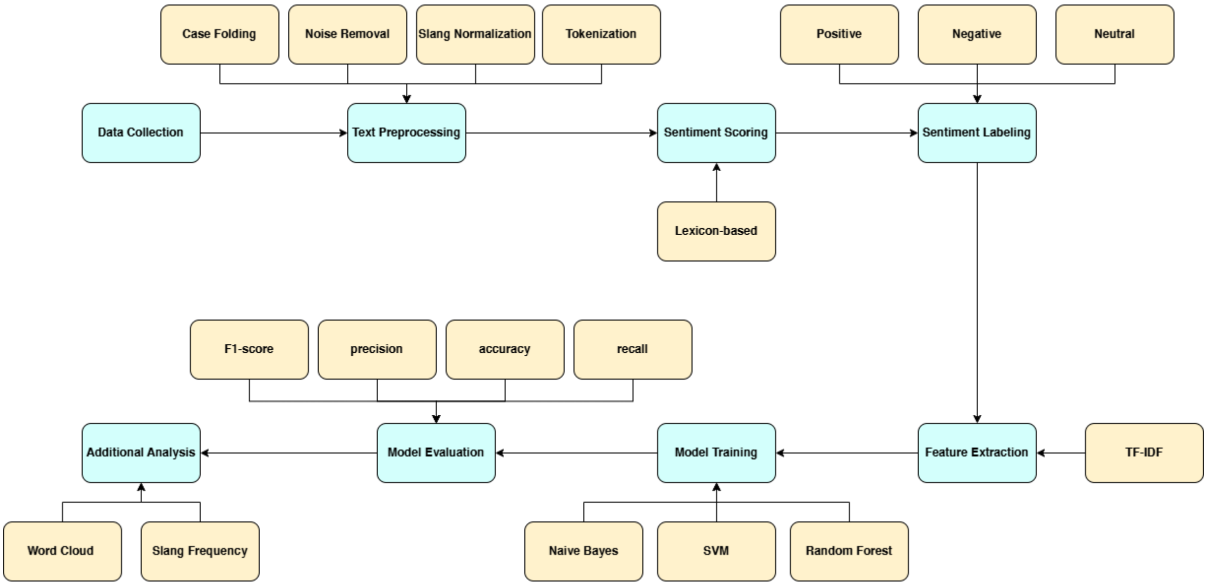

Figure 1. Research flow

The data collection and preprocessing phase began with acquiring textual user reviews that prominently featured informal and conversational language, including slang terms. The preprocessing process started with case folding, where all text was converted to lowercase for uniformity. Following this, noise removal was conducted to eliminate punctuation, numbers, and special characters using regular expressions. A critical component of preprocessing was slang normalization, where slang words were replaced with their formal equivalents using a predefined slang dictionary. Table 1 presents the slang dictionary used in this study, which played a key role in ensuring that informal terms were standardized. For example, words like "gak" (no) were normalized to "tidak", while "banget" (very) was replaced with "sangat". The normalized text was then tokenized into individual words, preparing it for further sentiment scoring and feature extraction.

Table 1. Slang Dictionary

Slang Word	Formal Equivalent (Bahasa Indonesia)	English Explanation
gak	Tidak	Means "no" or "not".
aja	Saja	Means "just" or "only".
banget	Sangat	Means "very" or "extremely".
gue	Saya	Informal word for "I" or "me".
lo	Kamu	Informal word for "you".
mantap	Bagus	Means "great" or "awesome".
nih	Ini	Means "this" or "here".
gratisan	Gratis	Refers to something "free of charge".
mahal	Mahal	Means "expensive" (same as formal word).
lambat	Lelet	Refers to something "slow".

To assign sentiment labels to the reviews, a lexicon-based sentiment scoring approach was employed. A sentiment lexicon (Table 2) was used to assign scores to individual words based on their sentiment polarity. Each word in a review was compared against the lexicon, and a cumulative sentiment score SSS was calculated using the following formula:

$$S = \sum_{i=1}^{n} \text{score}\ (w_i) \quad (1)$$

w_i represents each word in the review, and score (w_i) is the sentiment score assigned to that word. Reviews with a cumulative sentiment score greater than zero were labeled as Positive, those with scores less than zero were labeled as

Negative, and those with scores equal to zero were classified as Neutral. Table 2 provides a sample of the sentiment lexicon used in this study.

Table 2. Sentiment Lexicon

Word	Sentiment Score	Explanation
kecewa	-0.4	Indicates disappointment or dissatisfaction.
rugi	-1.0	Represents a significant negative sentiment, meaning "loss" or "unprofitable".
bagus	0.5	Represents a positive sentiment, meaning "good" or "nice".
mahal	-0.3	Indicates a slightly negative sentiment, meaning "expensive".
mantap	1.0	Represents a strong positive sentiment, meaning "excellent" or "awesome".
puas	0.9	Indicates high satisfaction or contentment.
lambat	-0.8	Represents a significant negative sentiment, meaning "slow".
susah	-0.6	Indicates difficulty or hardship, representing a negative sentiment.
promo	0.6	Represents a positive sentiment, meaning "promotion" or "special offer".
gagal	-0.5	Indicates failure or lack of success, representing a negative sentiment.

Following sentiment labeling, TF-IDF (Term Frequency-Inverse Document Frequency) was applied to transform the textual data into numerical features suitable for machine learning. TF-IDF assigns importance to words that occur frequently in a specific review but are less common across the entire dataset. The formula for computing the TF-IDF score for a term ttt in a document ddd is as follows:

$$\mathrm{TF-IDF}\ (\mathrm{t,d}) = \mathrm{TF}\ (\mathrm{t,d})\mathrm{x}\ \mathrm{IDf}\ (\mathrm{t}) \tag{2}$$

$\mathrm{TF}\ (\mathrm{t,d})$ is the term frequency, and $\mathrm{IDf}\ (\mathrm{t})$ is the inverse document frequency, calculated as:

$$\mathrm{IDF(t)} = \log\left(\frac{\mathrm{N}}{\mathrm{n_t}}\right) \tag{3}$$

N represents the total number of documents, and $\mathrm{n_t}$ is the number of documents containing the term t. TF-IDF effectively identifies terms that contribute most significantly to sentiment classification by balancing term frequency and document importance.

The processed features were then used to train three machine learning models: Naive Bayes, Support Vector Machine (SVM), and Random Forest. Naive Bayes, a probabilistic classifier, was selected as a baseline due to its efficiency in text classification. SVM was employed for its ability to handle high-dimensional data and identify the optimal hyperplane for separating sentiment classes. Random Forest, an ensemble learning technique, was chosen for its robustness and ability to reduce overfitting through the aggregation of multiple decision trees. The dataset was split into training and testing sets using an 80-20 split, ensuring that 80% of the data was used for model training and 20% for testing. Model performance was evaluated using standard metrics, including accuracy, precision, recall, and F1-score, to measure the models' effectiveness. Cross-validation was conducted to ensure that the models generalized well to unseen data.

Finally, additional analyses were performed to gain further insights into user feedback. A word cloud visualization was generated to highlight the most frequently occurring words in the reviews, providing a clear view of key themes. Terms like “sinyal,” “jaringan,” and “mahal” appeared prominently, indicating common issues related to network performance and pricing. Additionally, slang frequency analysis was conducted to identify the prevalence of slang terms within the dataset, emphasizing the necessity of normalization in sentiment analysis.

This methodological framework successfully integrates lexicon-based sentiment scoring, advanced text preprocessing, and machine learning techniques to address the challenges posed by informal and slang-enriched text. By combining these approaches, the study ensures accurate sentiment classification while providing actionable insights into user concerns and feedback.

4. Results and Discussion

4.1. Results

The results of the sentiment analysis are presented based on the classification of user reviews into Positive, Negative, and Neutral sentiments. The distribution of these sentiment categories is shown in table 3. Neutral sentiment accounted for the largest proportion of reviews, representing 51% of the total dataset. Negative sentiment followed with 28%, while Positive sentiment made up 21%. This distribution indicates that a majority of users provided factual or descriptive feedback without strong emotional undertones. However, a significant portion of the reviews expressed dissatisfaction, as reflected in the Negative sentiment category.

Table 3. The Distribution of Sentiment Categories

Sentiment	Count
Neutral	5090
Negative	2791
Positive	2119

Figure 2 visually illustrates the imbalance between the sentiment classes, emphasizing the dominance of Neutral feedback and the notable presence of Negative sentiment.

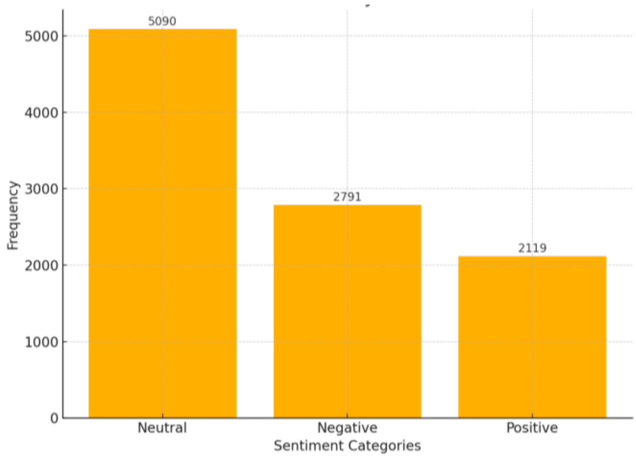

Figure 2. Sentiment Classification

To better understand the distribution of sentiment intensity, sentiment scores were calculated using a lexicon-based approach. Table 4 provides the frequency distribution of sentiment scores, and this is further visualized in Figure 2. The results show that most sentiment scores fall within the range of -1 to 1, with a noticeable peak at score 0, corresponding to Neutral sentiment. Negative sentiment scores ranged from -1 to -4, indicating varying levels of dissatisfaction, from mild frustrations to severe complaints. In contrast, positive sentiment scores clustered between 0.5 and 2, though they occurred much less frequently than negative scores. The histogram in figure 3 highlights this trend, showing that while the Neutral sentiment is dominant, negative reviews are spread across a broader range, reflecting the intensity of user dissatisfaction.

Table 4. The Frequency Distribution of Sentiment Scores

Score Range	Count
(-0.22, 0.41]	5495
(0.41, 1.04]	1988
(-0.85, -0.22]	1775
(-1.48, -0.85]	517
(-2.11, -1.48]	154

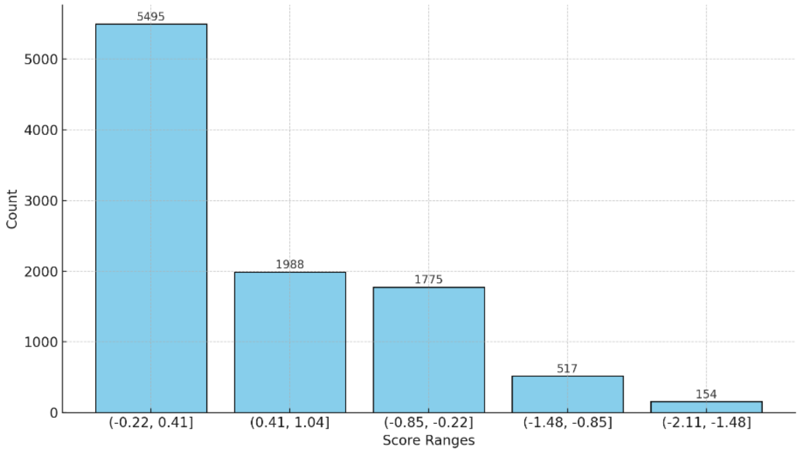

Figure 3. Score range distribution

A notable feature of the dataset is the frequent use of slang words, which reflects the informal and conversational tone of user feedback. Table 5 lists the most frequently occurring slang terms, with “gak” (meaning “no”) appearing 840 times, followed by “aja” (meaning “only”) with 700 occurrences and “banget” (meaning “very”) with 611 occurrences. Other frequently used slang terms include “nih” (150 occurrences) and “gue” (100 occurrences). These findings are visualized in figure 4, which shows the top slang terms in a horizontal bar chart. The predominance of informal expressions highlights the importance of effective text preprocessing, particularly slang normalization, to ensure accurate sentiment classification.

Table 5. Most Frequently Occurring Slang Terms

Slang Word	Count	Explanation
gak	829	Informal word for "no" or "not" (tidak).
aja	681	Informal word for "just" or "only" (saja).
banget	586	Informal word for "very" or "extremely" (sangat).
nih	135	Informal word for "this" or "here" (ini).
gue	52	Informal pronoun for "I" or "me" (saya).

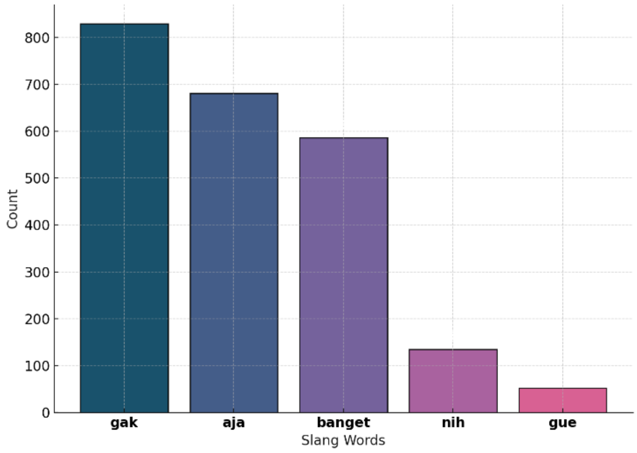

Figure 4. Frequency of slang words

A word cloud was generated to summarize the most frequently occurring words within the dataset, and the results are displayed in figure 5. Key terms such as “sinyal” (signal), “jaringan” (network), and “pulsa” (credit) emerged as the most prominent words, highlighting common themes within user reviews. Negative terms like “mahal” (expensive) and “lambat” (slow) also appeared frequently, reinforcing the findings of the sentiment classification. Positive words, such as “bagus” (good), “mantap” (great), and “baik” (excellent), were present but occurred less frequently. The word

ISSN 2723-6471

cloud provides a clear visualization of the dominant topics discussed by users, with a strong emphasis on network performance issues, service pricing, and overall reliability.

Figure 5. Word cloud result

As shown in table 6, slang normalization plays a crucial role in improving the performance of sentiment classification models by standardizing informal language into its formal equivalents. Without normalization, models struggle to interpret slang variations, leading to lower accuracy and misclassification of sentiments. The impact is most evident in Naive Bayes, which relies on word frequency and likelihood estimations. Without normalization, it achieves only 65% accuracy, while Support Vector Machine (SVM) and Random Forest perform relatively better at 88% and 89%, respectively. However, when slang normalization is applied, all models show significant improvements, with Naive Bayes reaching 73% accuracy and both SVM and Random Forest achieving 96%.

Table 6. Performance Comparison of Machine Learning Models with and without Slang Normalization

Model	Without Normalization				With Normalization			
	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
Naive Bayes	65%	68%	72%	69%	73%	75%	81%	74%
SVM	88%	90%	85%	87%	96%	97%	92%	94%
Random Forest	89%	91%	87%	89%	96%	96%	92%	94%

Beyond accuracy, normalization enhances precision, recall, and F1-score. Without normalization, models exhibit lower precision, particularly in detecting Negative sentiment, due to inconsistent slang usage. Recall is also affected, as models struggle to recognize sentiment-heavy reviews containing informal language. However, with normalization, both precision and recall improve, particularly in Naive Bayes, where recall increases from 72% to 81%, leading to better overall classification reliability. This improvement highlights the necessity of normalizing slang words to ensure that sentiment classifiers accurately interpret user opinions.

The significance of slang normalization lies in its ability to handle informal texts, reduce data sparsity, and enhance model interpretability. Many user-generated reviews contain conversational language that, when unprocessed, introduces inconsistencies in sentiment classification. Normalizing slang ensures that models train on a cohesive dataset, allowing them to generalize better and classify sentiments more accurately. Additionally, aligning machine learning models with lexicon-based scoring methods ensures consistency in sentiment prediction.

Overall, the results demonstrate that slang normalization is essential for improving sentiment classification performance. The most notable improvements occur in Naive Bayes, which benefits from an 8% accuracy boost, while SVM and Random Forest reach peak accuracy at 96%. These findings emphasize the importance of preprocessing in text classification, particularly when dealing with slang-rich content. Future research could explore deep learning models like BERT or LSTM, which may offer even better sentiment classification by capturing the contextual meaning of slang words without explicit normalization.

4.2. Discussion

The findings from the sentiment analysis provide valuable insights into user feedback and the challenges of analyzing informal, slang-enriched text. The dominance of Neutral sentiment (51%) suggests that many reviews are descriptive or factual without strong emotional tones. However, the significant presence of Negative sentiment (28%) highlights

ISSN 2723-6471

areas of user dissatisfaction, with sentiment scores revealing a broad range of intensity from mild complaints to severe frustrations. This pattern underscores the importance of identifying and addressing key pain points to improve customer experience.

One of the biggest challenges in sentiment analysis is the frequent use of slang words in user reviews. Common terms such as "gak," "aja," and "banget" reflect the conversational nature of user feedback, making it difficult for machine learning models to accurately interpret sentiment without proper preprocessing. Without slang normalization, sentiment classification models struggle to correctly process these informal expressions, leading to errors in classification. This study demonstrates that integrating slang normalization significantly improves the accuracy and reliability of sentiment analysis.

A comparative evaluation of model performance with and without slang normalization highlights its crucial role in improving classification accuracy, precision, recall, and F1-score. Without normalization, Naive Bayes performed the worst, achieving only 65% accuracy, while SVM and Random Forest showed moderate accuracy at 88% and 89%, respectively. After applying slang normalization, all models showed significant improvements—Naive Bayes improved to 73% accuracy, while SVM and Random Forest both reached 96% accuracy. Additionally, precision, recall, and F1-score increased across all models, demonstrating that normalization allows machine learning classifiers to interpret informal language more effectively and reduce misclassification errors.

Further analysis using word cloud visualization reveals key themes in user sentiment. Recurring terms such as "sinyal" (signal), "jaringan" (network), and "pulsa" (credit) indicate that users are primarily concerned with network reliability and service affordability. The frequent occurrence of “mahal” (expensive) and “lambat” (slow) reinforces that pricing and service performance are major sources of dissatisfaction. These findings suggest that users are particularly frustrated with network quality, service speed, and pricing strategies, contributing significantly to the volume of negative sentiment. Addressing these issues through network improvements and competitive pricing adjustments could reduce dissatisfaction and enhance user retention. On the other hand, positive sentiment was expressed far less frequently, as indicated by the low occurrence of words such as “bagus,” “mantap,” and “baik”. This suggests that strongly positive experiences are either less common or less frequently shared by users.

The evaluation of machine learning models further highlights the importance of selecting appropriate algorithms for sentiment classification, especially when dealing with slang-heavy text. Naive Bayes, while providing a useful baseline, struggled with 73% accuracy (even with normalization) and had difficulty in correctly identifying Negative sentiment, likely due to its reliance on simplistic probabilistic assumptions that fail to capture the complexity of informal expressions. In contrast, Support Vector Machine (SVM) and Random Forest significantly outperformed Naive Bayes, achieving 96% accuracy with normalization, while also maintaining high precision, recall, and F1-scores across sentiment categories. The superior performance of these models is attributed to their ability to handle high-dimensional text data and complex linguistic patterns, making them particularly effective in analyzing slang-enriched user reviews.

These findings underscore the importance of effective text preprocessing, particularly slang normalization, in improving sentiment classification accuracy. The strong presence of Negative sentiment, along with recurring complaints about network reliability, pricing, and service performance, highlights specific areas that service providers should focus on to enhance customer satisfaction. By leveraging these insights, companies can optimize service quality, refine pricing strategies, and improve user experience based on data-driven analysis.

Future research can build on these findings by incorporating deep learning models such as LSTM and BERT, which can better capture contextual relationships within text data and further improve sentiment classification performance. Additionally, unsupervised topic modeling techniques, such as Latent Dirichlet Allocation (LDA), could help identify specific subtopics driving user sentiment, providing a more granular understanding of customer concerns. These advanced approaches would enable a more nuanced and comprehensive analysis of user feedback, uncovering patterns that may not be evident through traditional sentiment classification methods.

ISSN 2723-6471

5. Conclusion

This study presents a comprehensive sentiment analysis of slang-enriched user reviews using lexicon-based sentiment scoring and machine learning approaches, providing key insights into user experience, recurring concerns, and the effectiveness of sentiment classification models in handling informal textual data. The results reveal that Neutral sentiment dominates (51%), while Negative sentiment (28%) highlights significant dissatisfaction, particularly regarding network quality, service speed, and pricing, as reflected in frequently mentioned terms such as "sinyal" (signal), "jaringan" (network), and "mahal" (expensive). The sentiment score distribution further emphasizes this trend, with negative sentiment exhibiting a wider intensity range compared to positive sentiment. A major challenge addressed in this study is the frequent use of slang words, which complicates sentiment classification; however, the successful implementation of slang normalization significantly improved model accuracy. Support Vector Machine (SVM) and Random Forest outperformed Naive Bayes, achieving 96% accuracy compared to 73% for Naive Bayes, demonstrating their robustness in handling high-dimensional, slang-enriched text. Furthermore, slang normalization played a crucial role in improving model performance, confirming its importance in enhancing text preprocessing and classification accuracy. These findings highlight that network reliability, affordability, and service speed are key drivers of user dissatisfaction, presenting opportunities for service providers to enhance customer experience. Addressing these issues could significantly reduce negative sentiment and improve overall user satisfaction. In conclusion, this study underscores the necessity of effective text preprocessing, particularly slang normalization, for improving sentiment classification. Future research can explore deep learning techniques such as BERT or LSTM to capture contextual relationships within text data and further refine sentiment classification, while topic modeling methods can be employed to uncover deeper themes in user feedback, enabling data-driven improvements in service quality and customer experience.

6. Declarations

6.1. Author Contributions

Conceptualization: P.A.P., B., and I.T.; Methodology: B.; Software: P.A.P.; Validation: P.A.P., B., and I.T.; Formal Analysis: P.A.P., B., and I.T.; Investigation: P.A.P.; Resources: B.; Data Curation: B.; Writing Original Draft Preparation: P.A.P., B., and I.T.; Writing Review and Editing: B., P.A.P., and I.T.; Visualization: P.A.P.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

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

References

[1] C. Kumaresan and P. Thangaraju, "Sentiment Analysis in Multiple Languages: A Review of Current Approaches and Challenges," *REST Journal on Data Analytics and Artificial Intelligence*, vol. 2, no. 1, pp. 1-12, 2023. DOI: 10.46632/jdaai/2/1/2.

[2] K. Sharma, D. Jain, A. Jain, Y. S. Rathore, and S. Agarwal, "Sentimental Analysis based on Machine Learning Technique," in *2024 3rd International Conference for Innovation in Technology (INOCON)*, vol. 3, no. 1, pp. 1–7, 2024. DOI: 10.1109/INOCON60754.2024.10511973.

[3] F. Xu, X. Zhang, Z. Xin, and A. Yang, "Investigation on the Chinese Text Sentiment Analysis Based on Convolutional Neural Networks in Deep Learning," *Computers, Materials & Continua*, vol. 62, no. 2, pp. 1-12, 2019. DOI: 10.32604/cmc.2019.05375.

[4] S. K. Papia, M. A. Khan, T. Habib, M. Rahman, and M. N. Islam, "DistilRoBiLSTMFuse: an efficient hybrid deep learning approach for sentiment analysis," *PeerJ Comput. Sci.*, vol. 10, no. 1, pp. 1-16, 2024. DOI: 10.7717/peerj-cs.2349.

[5] B. Gedif, A. Alemu, Y. Assefa, and S. Nibret, "Design Amharic Text Sentiment Analysis Model Using Machine Learning Techniques," in *2023 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, vol. 2023, no. 1, pp. 150–154, 2023. DOI: 10.1109/ICT4DA59526.2023.10302239.

[6] M. Alfreihat, O. Almousa, Y. M. Tashtoush, A. AlSobeh, K. Mansour, and H. Migdady, "Emo-SL Framework: Emoji Sentiment Lexicon Using Text-Based Features and Machine Learning for Sentiment Analysis," *IEEE Access*, vol. 12, no. 1, pp. 81793–81812, 2024. DOI: 10.1109/ACCESS.2024.3382836.

[7] A. Ashir, "A Generalized Method for Sentiment Analysis across Different Sources," *Applied Computational Intelligence and Soft Computing*, vol. 2021, no. 1, pp. 1–8, 2021. DOI: 10.1155/2021/2529984.

[8] M. Abdul Harsath, A. Arshad, H. R., and R. Srikanth, "Improving Sentiment Analysis of Social Media Captions Through Advancements in NLP," in *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, vol. 2024, no. 1, pp. 732–736, 2024. DOI: 10.1109/ICSCSS60660.2024.10624998.

[9] A. Geet, R. Yadav, S. K. Alaria, S. K. Gupta, and J. Bhutto, "Numerical Simulation and Design of Sentiment Analysis of Twitter Posts Using Hybrid Machine Learning Approach," in *Proceedings of the 5th International Conference on Information Management & Machine Intelligence*, vol. 2023, no. 1, pp. 1-8, 2023. DOI: 10.1145/3647444.3652443.

[10] H. Raza, M. Faizan, A. Hamza, A. Mushtaq, and N. Akhtar, "Scientific Text Sentiment Analysis using Machine Learning Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 1, pp. 12-22, 2019. DOI: 10.14569/ijacsa.2019.0101222.

[11] M. Alfreihat, O. Almousa, Y. M. Tashtoush, A. AlSobeh, K. Mansour, and H. Migdady, "Emo-SL Framework: Emoji Sentiment Lexicon Using Text-Based Features and Machine Learning for Sentiment Analysis," *IEEE Access*, vol. 12, no. 1, pp. 81793–81812, 2024. DOI: 10.1109/ACCESS.2024.3382836.

[12] J. Hartmann, J. Huppertz, C. Schamp, and M. Heitmann, "Comparing automated text classification methods," *Int. J. Res. Mark.*, vol. 36, no. 1, pp 1-13, 2019. DOI: 10.1016/j.ijresmar.2018.09.009.

[13] M. Arumugam, S. S. R., and C. Jayanthi, "Machine Learning for Sentiment Analysis Utilizing Social Media," in *2023 2nd Int. Conf. Edge Comput. Appl. (ICECAA)*, vol. 2, no. 1, pp. 523–530, 2023. DOI: 10.1109/ICECAA58104.2023.10212135.

[14] M. Deepak I., "Sentiment Analysis Techniques in Recent Works Using GRA Methodology," *J. Bus. Anal. Bull.*, vol. 1, no. 3, pp. 1-13, 2023. DOI: 10.46632/jbab/1/3/8.

[15] V. M. Abia and E. H. Johnson, "Sentiment Analysis Techniques: A Comparative Study of Logistic Regression, Random Forest, and Naive Bayes on General English and Nigerian Texts," *J. Eng. Res. Rep.*, vol. 26, no. 1, pp. 1-12, 2024. DOI: 10.9734/jerr/2024/v26i91268.

[16] E. Shah Aly and D. Haar, "Slang-Based Text Sentiment Analysis in Instagram," in *Advances in Intelligent Systems and Computing*, Springer, vol. 2019, no. 1, pp. 321–329, 2019. DOI: 10.1007/978-981-32-9343-4_25.

[17] V. U., V. M. V., and P. R. M., "Decoding Sentiments: Virtue or Vice through Multilingual Paragraph Analysis," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 2024, no. 1, pp. 1-14, 2024. DOI: 10.48175/ijarsct-15312.

[18] S. Soumya and V. PramodK., "Sentiment Analysis of Malayalam Tweets Using Machine Learning Techniques," *ICT Express*, vol. 6, no. 1, pp. 300–305, 2020. DOI: 10.1016/j.icte.2020.04.003.

[19] E. Fergan and T. M. Tashu, "Course Review Sentiment Analysis: A Comparative Study of Machine Learning and Deep Learning Methods," in *2023 10th Int. Conf. Behav. Soc. Comput. (BESC)*, vol. 10, no. 1, pp. 1–7, 2023. DOI: 10.1109/BESC59560.2023.10386515.

[20] M. S. Başarslan and F. Kayaalp, "Sentiment Analysis on Social Media Reviews Datasets with Deep Learning Approach," *SAUCIS*, vol. 4, no. 1, pp. 35–49, 2021. DOI: 10.35377/SAUCIS.04.01.833026.

[21] M. M. J. Soumik, S. S. M. Farhavi, F. Eva, T. Sinha, and M. S. Alam, "Employing Machine Learning Techniques on Sentiment Analysis of Google Play Store Bangla Reviews," in *2019 22nd Int. Conf. Comput. Inf. Technol. (ICCIT)*, vol. 22, no. 1, pp. 1–5, 2019. DOI: 10.1109/ICCIT48885.2019.9038348.

[22] L. Yao, J. Li, H. Alam, and O. Melnokov, “An Evaluation of Tweet Sentiment Classification Methods,” in *2020 Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, vol. 2020, no. 1, pp. 298–303, 2020. DOI: 10.1109/CSCI51800.2020.00057.

[23] G. Kathuria, K. Singh, S. P. Majgahe, and K. Kumar, “Machine Learning Based Sentiment Analyzer,” in *2024 2nd Int. Conf. Device Intell. Comput. Commun. Technol. (DICCT)*, vol. 2024, no. 1, pp. 88–93, 2024. DOI: 10.1109/DICCT61038.2024.10532857.

[24] H. D. Abubakar and M. Umar, “Sentiment Classification: Review of Text Vectorization Methods: Bag of Words, TF-IDF, Word2vec and Doc2vec,” *SLU J. Sci. Technol.*, vol. 4, no. 1, pp. 1-13, 2022. DOI: 10.56471/slujst.v4i.266.

[25] T. Tesfa, B. Belete, S. Abera, S. K. Mohapatra, and T. K. Das, “Aspect-Based Sentiment Analysis on Amharic Text for Evaluating Ethio-Telecom Services,” in *2024 2nd Int. Conf. Emerg. Trends Inf. Technol. Eng. (ICETITE)*, vol. 2, no. 1, pp. 1–6, 2024. DOI: 10.1109/ic-ETITE58242.2024.10493748.