SinkrOn
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15644
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Emotion-Based Multi-Class Sentiment Analysis of FirstMedia Customers Reviews Using SVM with Kernel Comparison

Bagus Kustiono Ongko[1], Anik Vega Vitianingsih[2]*, Dwi Cahyono[3], Anastasia Lidya Maukar[4], Seftin Fitri Ana Wati [5]
[1,2,3]Informatics Department, Universitas Dr. Soetomo, Surabaya, Indonesia
[4]Industrial Engineering Department, President University, Bekasi, Indonesia
[5]Information System Department, Universitas Pembangunan Nasional "Veteran" Jawa Timur, Indonesia
[1]baguskustiono@gmail.com, [2]vega@unitomo.ac.id, [3]dwik@unitomo.ac.id, [4]almaukar@president.ac.id,
[5]seftin.fitri.si@upnjatim.ac.id

Submitted : Nov 30, 2025 | **Accepted** : Dec 29, 2025 | **Published** : Jan 4, 2026

Abstract: The advancement of digital technology has made users increasingly reliant on online services, with user reviews serving as an essential resource for evaluating the quality of service provided by companies such as FirstMedia. However, these valuable data have not undergone comprehensive analysis to assess users' emotional responses. This study aims to classify FirstMedia customers' emotions into four categories (joy, sadness, anger, and neutral) and to evaluate the Support Vector Machine (SVM) method using four different kernel functions. Most existing studies primarily focus on polarity-based sentiment analysis and do not explicitly examine multi-emotion classification or kernel comparison in machine learning models. A total of 4,001 reviews were collected through web scraping from the Google Play Store and the X app and processed through several preprocessing steps. Emotion classification was conducted using the NRC Indonesian Emotion Lexicon, while word significance was determined using TF-IDF weighting. After preprocessing, 3,069 labeled reviews were retained and distributed as 1,065 neutral, 748 anger, 692 joy, and 564 sadness reviews, which were used for emotion classification. Model performance was evaluated using a hold-out validation scheme with an 80:20 train-test split and assessed through a confusion matrix. To address class imbalance, undersampling was applied, resulting in a balanced dataset for model training. The evaluation results show that the Linear kernel achieved the highest performance, with an accuracy of 82.63%, precision of 82.86%, recall of 82.63%, and an F1-score of 82.60%, outperforming the Gaussian, Polynomial, and Sigmoid kernels. This study demonstrates that multi-emotion sentiment analysis provides a more comprehensive understanding of user perceptions beyond conventional sentiment polarity, thereby supporting more informed evaluations of digital service quality.

Keywords: Sentiment Analysis; FirstMedia Sentiment Analysis; Support Vector Machine; Indonesian NRC Emotion Lexicon; Machine Learning.

INTRODUCTION

Public demand for internet services has increased significantly in line with advancements in digital technology (Bumbungan, 2025). Internet services have become essential for fulfilling various public needs in entertainment, work, education, and communication (Iswaratama, 2024). FirstMedia, as a provider of internet and cable television services in Indonesia, has become one of the preferred choices for meeting these connectivity requirements (Agung & Suryati, 2024). However, as the use of FirstMedia's digital services continues to grow, a wide range of user reviews has emerged across social media platforms such as the Google Play Store and X. These reviews vary considerably, ranging from positive experiences to complaints about slow connection quality, frequent disruptions, and service provider responses perceived as inadequate.

As the number of incoming reviews increases, manual analysis methods become increasingly inefficient in providing a clear representation of FirstMedia's service quality (Surbakti PP & Ginting AA, 2022). To address

*Anik Vega Vitianingsih

this gap, a system capable of filtering, grouping, and analysing the emotions embedded in each review is required. Previous studies have generally been limited to conventional sentiment analysis, which classifies reviews as positive, negative, or neutral without considering deeper emotional dimensions (Kadek et al., 2025). This limitation restricts a more comprehensive understanding of customer satisfaction and dissatisfaction. Limited studies have specifically explored multi-emotion sentiment analysis on FirstMedia user reviews, particularly those that classify emotions such as anger, joy, sadness, and neutral within a unified analytical framework. Therefore, the system offered in this study is anticipated to aid prospective customers in choosing an internet service provider and furnish vital information for FirstMedia's management to enhance service quality.

Previous studies (Muhammad Fikri et al., 2022) have used sentiment analysis with positive, negative, and neutral classifications to assess user experiences. In that study, the analysis of 8,089 tweets reviewing FirstMedia's services using the Ensemble Stacking method achieved an accuracy of 92.56%. A more recent study (Putri et al., 2024) analysing 696 sentiment-related tweets concerning the influence of brand ambassadors using the KNN and SVM methods found that KNN achieved an average accuracy of 71%, while SVM produced a superior average accuracy of 83%. Another study (Naufal et al., 2023) examining the performance of SVM, RF, and NB in cyberbullying classification also demonstrated that the SVM method achieved the highest accuracy of 83%, outperforming NB (74%) and RF (82%).

Based on these trends, this study proposes the use of the SVM method by employing four kernel functions (Praghakusma & Charibaldi, 2021)—Linear, Polynomial, Gaussian, and Sigmoid—to evaluate the capability of each kernel in managing high-dimensional datasets and non-linear classification tasks. The SVM method is well known for its stability and reliable performance, particularly when dealing with imbalanced data, as well as for its ability to classify multiple emotional categories accurately, making it effective in representing user experiences (Baliputra et al., 2025). This procedure incorporates the SVM algorithm alongside text preprocessing, TF-IDF weighting, and lexicon-based labelling (Jaya & Zeniarja, 2024). In addition, model performance is evaluated using confusion matrices.

LITERATURE REVIEW

In the study by (March Vircan Karuna, 2023), customer sentiment toward FirstMedia's Wi-Fi service was analysed using tweets from the @FirstMediaCares account. The study initially collected 845 tweets, of which only 310 were retained after the data-cleaning process. Sentiment classification was performed using the Naïve Bayes (NB) method in RapidMiner, resulting in 88% negative sentiment and 12% positive sentiment. These findings indicate a strong dominance of negative perceptions regarding FirstMedia's Wi-Fi service. However, the classification was limited to only two categories, thus failing to capture a broader range of emotional responses.

A separate study by (Prasetio et al., 2021), examined consumer perceptions of the MyIndiHome service by comparing the effectiveness of SVM and NB algorithms. Their results showed that SVM achieved a higher accuracy of 86.5% compared to NB's 84.6% when classifying user sentiment. Nonetheless, the study focused solely on general sentiment polarity and did not address more nuanced emotional categories or the specific contexts underlying user concerns.

Another study (Damayanti et al., 2024), analysed sentiment in product reviews of the Alfagift application using a dataset of 1,000 Google Play Store reviews. This study compared LSTM and SVM models for lexicon-based labelling and evaluation using a confusion matrix. The results demonstrated that SVM achieved the highest accuracy at 83.5%, slightly outperforming LSTM, which achieved 82%. The classification produced 801 positive reviews, 77 negative reviews, and 122 neutral reviews. Although the performance was satisfactory, the study focused solely on sentiment polarity and did not provide a comprehensive analysis of emotional dimensions.

In study (Garc et al., 2024), which explains the relationship between sentiment analysis and emotion recognition, the objectives are different but interrelated. Traditional sentiment analysis focuses on determining the polarity of opinions, such as positive, negative, or neutral, making it simpler and more effective for providing a general overview of public perception. However, this approach is often unable to capture deeper affective nuances. In contrast, emotion classification aims to identify specific emotions such as anger, happiness, and sadness, thereby providing a richer and more contextual understanding of user expressions. In the context of modern text analysis, emotion-based sentiment analysis can be considered superior because it not only determines the direction of sentiment but also explains the underlying type of emotion, thus serving as part of emotion recognition(Garc et al., 2024).

Study (Venugopal et al., 2024) show that most sentiment analysis research still places too much emphasis on improving accuracy, while conceptual aspects such as the difference between polarity classification and emotion classification, the limitations of lexicon-based labelling, and domain bias are often overlooked. This study confirms that models with high accuracy do not necessarily produce fair predictions, especially when the training data is dominated by a particular domain and lexicon-based labelling fails to capture emotional context and linguistic nuances (Venugopal et al., 2024). These findings reinforce the need to shift from polarity-based sentiment analysis to more affectively rich approaches, such as emotion-based sentiment analysis.

*Anik Vega Vitianingsih

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15644
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Based on these prior studies, SVM consistently demonstrates superior performance compared with other algorithms such as NB, RF, LSTM, and KNN, making it a highly reliable and suitable choice for sentiment analysis. However, these studies predominantly focus on positive–neutral–negative polarity classification and do not capture more specific emotional dimensions. However, the emphasis on accuracy alone has led many studies to remain limited to polarity-based sentiment classification, which restricts the ability to capture more nuanced emotional states. Recent findings suggest that emotion-based sentiment analysis offers richer interpretability than polarity classification, particularly in service evaluation contexts. Nevertheless, lexicon-based emotion labeling approaches are known to suffer from contextual insensitivity and domain dependency, potentially introducing bias when applied to specific service domains such as internet service providers. In addition, sentiment models trained on general datasets may exhibit domain bias (Tan et al., 2023), reducing their effectiveness when applied to platform-specific user reviews. These limitations indicate the need for an integrated approach that combines lexicon-based emotion labeling with machine learning models and evaluates kernel performance to better handle high-dimensional and domain-specific textual data.

METHOD

This analysis aims to evaluate user perceptions of FirstMedia's internet services through emotion-based sentiment analysis. The study was conducted to identify emotional patterns within user reviews and to assess the effectiveness of the SVM method for classification. The analysis commenced with the acquisition of user review data via web scraping from the Google Play Store and X App platforms, followed by the preparation of the data for subsequent processing. The gathered reviews underwent multiple preprocessing stages, including tokenization, data cleansing, case normalization, stopword elimination, and stemming. Subsequently, sentiment labeling was performed with a lexicon-based methodology by correlating terms in the evaluations with the Indonesian NRC Emotion Lexicon, which classifies emotions into joy, anger, sadness, and neutral (Surya et al., 2023). It should be noted that the emotion labels were generated using lexicon-based scoring rather than manual human annotation. Therefore, the assigned labels represent proxy emotion labels and do not constitute ground-truth emotional states. The importance of each word in the text was subsequently evaluated by utilizing TF-IDF weighting on the labeled data. The system design utilized in this study is depicted in Figure 1.

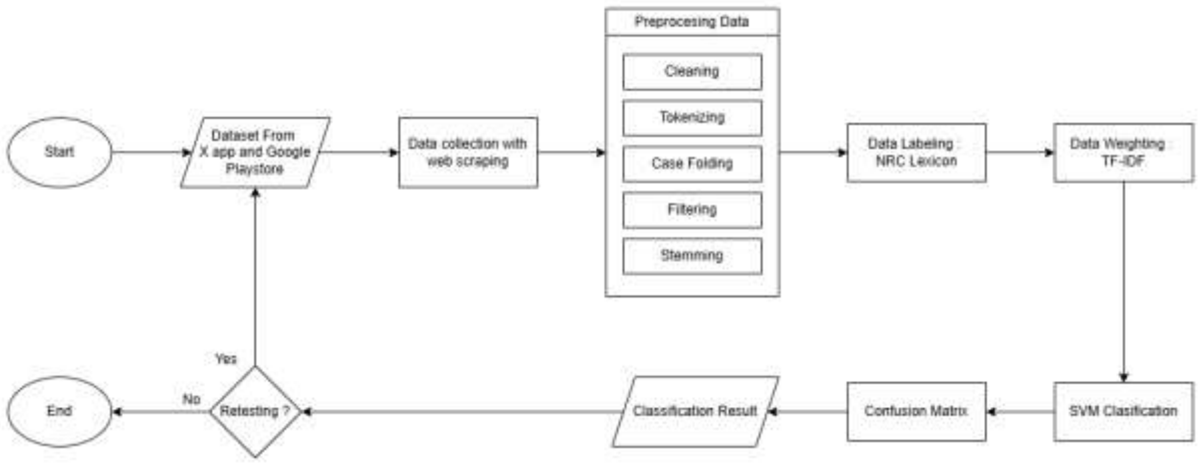


Fig 1. Design System Flow

The weighted data were analyzed using SVM, comparing kernels classification to categorise the reviews into the predetermined emotion groups. A confusion matrix was employed to evaluate the effectiveness of the model (Vanacore et al., 2024). Thus, the classification task in this study reflects emotion prediction based on lexicon-derived labels, and the evaluation focuses on comparing kernel performance rather than claiming definitive psychological emotion recognition.

Data Collection

This study utilises user review data from the X application and the Google Play Store regarding internet connection services collected between January 2024 and November 2025. The data were gathered using Python-based tools such as TweetHarvest and Google Play Scraper. The reviews analysed in this study consist of Indonesian-language user opinions related to FirstMedia's services, collected using the keywords #InternetFirstMedia or #FirstMedia.

*Anik Vega Vitianingsih

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Preprocesing

Data pre-processing is an essential stage in text analysis that readies raw service review data for optimal application in SVM models. (Putra Selian et al., 2024). This study employs pre-processing techniques, including cleaning, case folding, stemming, filtering, and tokenization. These strategies facilitate the generation of cleaner, more pertinent data, hence enhancing the accuracy and efficacy of sentiment analysis of the examined text (Palomino & Aider, 2022). The following are the fundamental steps employed in data pre-processing:

Step 1. Cleaning: The phase eliminates extraneous elements, including symbols, numerals, and URLs, retaining only pertinent words for sentiment analysis.

Step 2. Case Folding: The stage converts all text to lowercase for uniformity, treating terms like 'HAPPY' and 'happy' as the same.

Step 3. Tokenization: The stage disaggregates text into its minimal components, such as words or phrases, thereby facilitating the model's comprehension of its structure and enabling subsequent analysis.

Step 4. Filtration: The stage eliminates words that lack substantial informational value for the analysis, including common terms and stopwords that are extraneous to the sentiment context.

Step 5. Stemming: Reduces words to their basic or root form, for example, converting 'run' and 'running' to 'run', hence minimizing variations of words with identical meanings

Data Labelling

The data labelling process was automated using a lexicon-based approach, in which each word in a review was scored according to a sentiment dictionary (Simay Akar et al., 2024). The total score for each review was obtained by summing the sentiment values of each word. Classification was performed using NRC Lex's emotion analysis, which assesses four primary emotions in text: anger, joy, sadness, and neutral. If the text has the highest score for anger, joy, or sadness, the review is categorized by that emotion. If there is no dominant emotion or the score for that emotion is low, the review is categorized as neutral.

Emotion labeling using the NRC Indonesian Emotion Lexicon, it should be noted that the assigned emotion labels were derived from lexicon-based scoring rather than manual human annotation. Consequently, the resulting labels represent proxy emotion labels instead of ground-truth emotional states. This approach is commonly adopted in large-scale sentiment analysis studies due to its scalability and consistency, particularly when manual annotation is impractical. The use of proxy labels in this study aims to support comparative evaluation of classification models rather than to claim definitive psychological emotion recognition (Zad et al., 2021).

Word Weighting

The TF-IDF technique assesses a word's importance in a text by analyzing its frequency within a document and its scarcity across the entire document corpus, as shown in Equation (1) (Danyal et al., 2024). TF quantifies the frequency of a word within a given text, whereas IDF assigns higher values to words that are rare across multiple documents, indicating greater significance. The combination of TF and IDF yields a TF-IDF score that indicates the significance of a term within a document relative to the full corpus. Terms that are prevalent in a single document and infrequent in others will possess an elevated TF-IDF score (Rahmadani et al., 2024)

$$tf - idf_{t,d} = w_{tf\ t,d} \times idf_t \quad (1)$$

Split Data

The data-splitting phase in sentiment analysis plays a critical role, as it divides the dataset into two primary components: training data and testing data (Qi & Shabrina, 2023). This procedure ensures that a machine learning model can effectively learn underlying patterns before being assessed on previously unseen information. During the training phase, the model internalizes various relationships, structures, and features present in the dataset. It is then evaluated using the testing subset to determine its accuracy, generalization ability, and overall performance in predicting sentiment (Tan et al., 2023). Model evaluation was conducted using a single hold-out validation scheme with an 80:20 train–test split. To ensure reproducibility, a fixed random seed was applied during the splitting process. Although cross-validation and repeated runs may provide more robust performance estimates, this study adopted a hold-out strategy to maintain computational efficiency and to focus on kernel comparison.

SVM

SVM is a method for data classification that determines the best hyperplane to differentiate between two classes, as articulated in Equation (2) (Angelie Tania & Oetama, 2025). SVM optimizes the margin between the two classes, specifically the distance from the hyperplane to the nearest data points. SVM employ the Structural Risk Minimisation (SRM) concept to identify the hyperplane that minimizes generalization error and can be modified to address non-linear issues using linear, polynomial, Gaussian, and sigmoid kernel functions (Anggoro &

*Anik Vega Vitianingsih

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15644
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Permatasari, 2023). Linear kernels are used for linearly separable data.(Shamsi & Beheshti, 2025) Hyperparameters used Linear C = 1.0, RBF C = 1.0, gamma = scale, Polynomial C = 1.0, degree = 3, gamma = scale, Sigmoid C = 1.0, gamma = scale.

The application of these four kernels facilitates the examination of diverse decision thresholds and helps determine which kernel best captures the data's sentiment patterns (Fauzi et al., 2025).

Linear Kernel, used Equation (2) to calculate the similarity between two data vectors.

$$K(x_i, x) = x_i \, . \, x \quad (2)$$

Polynomial kernel, used Equation (3) for data separated by complex decision boundaries based on the degree of the polynomial.

$$K(x_i, x) = (x_i \cdot x)^d \quad (3)$$

Gaussian kernel, calculated using Equation (4) for similarity based on Euclidean distance.

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) \quad (4)$$

Sigmoid kernel, using Equation (5) for tanh activation function, similar to artificial neural networks for non-linear classification.

$$K(x_i, x) = \tanh\left(\sigma(x_i \, . x) + c\right) \quad (5)$$

Confusion Matrix

The efficacy of the SVM model is evaluated using a Confusion Matrix, which compares the model's predictions with the actual labels to assess classification accuracy and identify errors, including false positives (FP) and false negatives (FN) (Hendrastuty et al., 2021). This matrix enables the calculation of several critical evaluation metrics, namely accuracy, precision, recall, and F1-score (Helmud et al., 2024). Accuracy quantifies the model's overall classification effectiveness; precision evaluates the model's correctness in predicting the positive class; recall assesses the model's capacity to identify true positives; and the F1-score harmonizes precision and recall, which is crucial for evaluating class performance. These four metrics offer a thorough assessment of the model's classification accuracy. (Helmud et al., 2024).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$precision = \frac{TP}{TP+FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (8)$$

$$\text{F1-Score} = \frac{2 \times \text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (9)$$

RESULT

Data Colection

This study successfully collected 4001 user reviews on the Google Play Store and X App using web scraping. Data was obtained through keyword searches, such as "FirstMedia," "FirstMedia Internet," and other terms related to user experiences using First Media's internet services. The collected reviews covered key aspects, including internet speed, disconnections, and the package's pricing. Table 1 shows examples of reviews used in the sentiment analysis.

Table 1. Scraping Data Result from Google Play Store

Example review data from Google Play Store:
Duhh!!! Bikin kesel nih first Media, tbtb cs tlfn mau kasih free tambah Mbps selama 6 bulan, belum ada kelanjutan tapi internet ga bisa di pake. Modemnya on tapi ga ada koneksi Internetnya, komplen di aplikasi susahnya ampunnnn
Udah 5 tahun lebih langganan firstmedia dan selalu puas sama layanannya. Wifi stabil dan siaran tv jg bagusss... Kalo ada masalah, cepet ditanganinnya lewat app myfirstmedia ini. Tinggal restart modem, biasanya

*Anik Vega Vitianingsih

langsung udah bener lagi. Kalo masih eror, bisa langsung tlp call center / mention di twitter. Thank you firstmedia ❤️
Awal mula pakai ini internet lancar, harga juga bersahabat. Skrg harga naik ok gpp, tp internet ttp lemot, apalagi pas pagi-sore lemotnya setengah mati. Tolong di perbaiki

Example review data from X apps:
internet di rumah setiap 5 menit pasti off Kalau lagi zoom otomatis kelempar Padahal speednya bisa sampai 60mbps Dari kemarin ganti perangkat tidak ketemu solusinya. Tolong lah di respon sudah melalui myfirstmedia aja dicuekin
haduh min ini kenapa seharian lemot bgt download speed 50kbps??? Buka apps myfirstmedia dibilang ga ada gangguan alias koneksi ijo. Yg bener aja hei uid
Udah di restarr 3x hari ini min pake https://t.co/wN4JijE2d3 . Awal di restart mah lancar terus abis itu gak stabil lagi

Data Preprocesing

The initial phase of data processing is pre-processing, which encompasses numerous critical steps. Cleaning is conducted to eliminate extraneous items, including symbols, numbers, and URLs. Subsequently, filtering eliminates words that lack substantial informational value, such as stop words. Case folding transforms all text to lowercase to guarantee uniformity (Berutu et al., 2023). Tokenization divides the text into its basic components, such as words or sentences. Ultimately, stemming reduces words to their roots. All stages are delineated in Table 3.

Table 3. Preprocesing Step

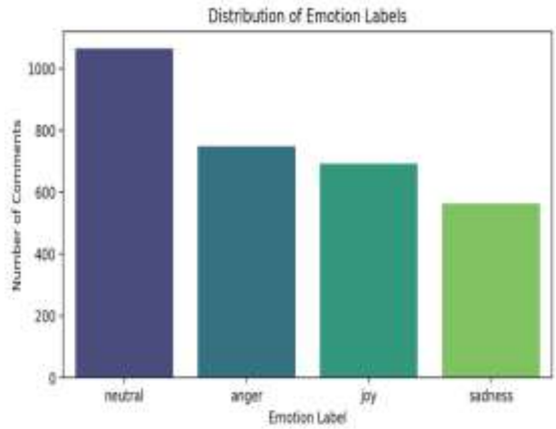
Pre-procesing step	Result
Dataset	*Koneksi internet gabisa daritadi pagi status di myfirstmedia juga ONLINE. Terakhir perbaikan 12 maret kemarin kok udah trouble lagi ya :v*
Cleaning	*Koneksi internet tidak dapat digunakan sejak pagi, status di MyFirstMedia juga ONLINE. Terakhir perbaikan pada 12 Maret, namun sudah terjadi gangguan lagi.*
Case Folding	*koneksi internet tidak dapat digunakan sejak pagi, status di myfirstmedia juga online. terakhir perbaikan pada 12 maret, namun sudah terjadi gangguan lagi.*
Tokenizing	*["koneksi", "internet", "tidak", "dapat", "digunakan", "sejak", "pagi", "status", "di", "myfirstmedia", "juga", "online", "terakhir", "perbaikan", "pada", "12", "maret", "namun", "sudah", "terjadi", "gangguan", "lagi"]*
Filtering	*["koneksi", "internet", "digunakan", "pagi", "status", "myfirstmedia", "online", "perbaikan", "maret", "terjadi", "gangguan"]*
Stemming	*["koneksi", "internet", "guna", "pagi", "status", "myfirstmedia", "online", "baik", "maret", "jadi", "ganggu"]*

Data Labeling

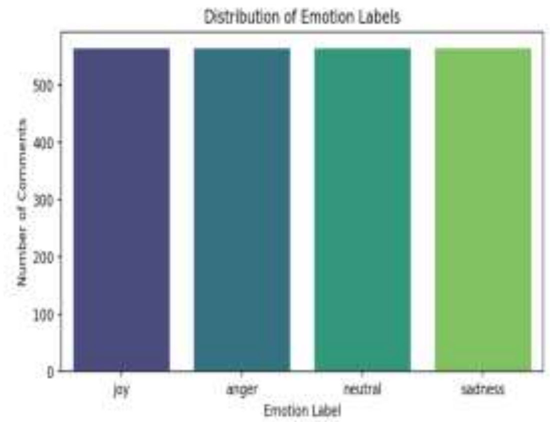
After applying lexicon-based emotion labeling using the NRC Indonesian Emotion Lexicon, a total of 3,069 labeled reviews were retained for further analysis. The labeled dataset showed an imbalanced distribution, consisting of 1,065 neutral reviews, 748 anger reviews, 692 joy reviews, and 564 sadness reviews. This imbalance indicates that neutral expressions were more dominant in user reviews, which may lead to biased classification outcomes if directly used for model training.

Therefore, undersampling was applied to reduce the majority class and produce a balanced dataset (Chang et al., 2021). After undersampling, each emotion class contained 564 reviews, resulting in a total of 2,256 reviews. This balanced dataset was subsequently used as input for the SVM classification stage. Need to remember, Undersampling may introduce sampling bias because some informative instances from the majority class are removed, which can reduce the representativeness of the dataset and affect model generalizability. The results of this data labelling process are shown in Figure 2(a) to (b).

*Anik Vega Vitianingsih



(a) Data Labelling Result



(b) Data Labelling after Undersampling

Word Weighting

Upon completion of the pre-processing phase of the sentiment analysis, the subsequent step is to implement TF-IDF weighting. The TF-IDF computation is presented in Table 4.

Table 4. Word Weighting Datasets

Word	Weighting TF-IDF
koneksi	0,0434
internet	0,0434
guna	0,0434
pagi	0,0434
status	0,0434
myfirstmedia	0,0434
online	0,0434
baik	0,0434
maret	0,0434
jadi	0,0434
ganggu	0,0434

The final result of this aggregation is visualized as a word cloud, with the most weighty words appearing more prominently. This visualization aims to facilitate the analysis of sentiment characteristics in user reviews. Figure 3 displays a word cloud from Google Play Store and X app reviews.

Fig 3. Word Cloud Datasets

Model Evaluation

The classification model's performance is assessed to determine its ability to categorize data accurately. Table 5 displays the assessment outcomes of the SVM methodology.

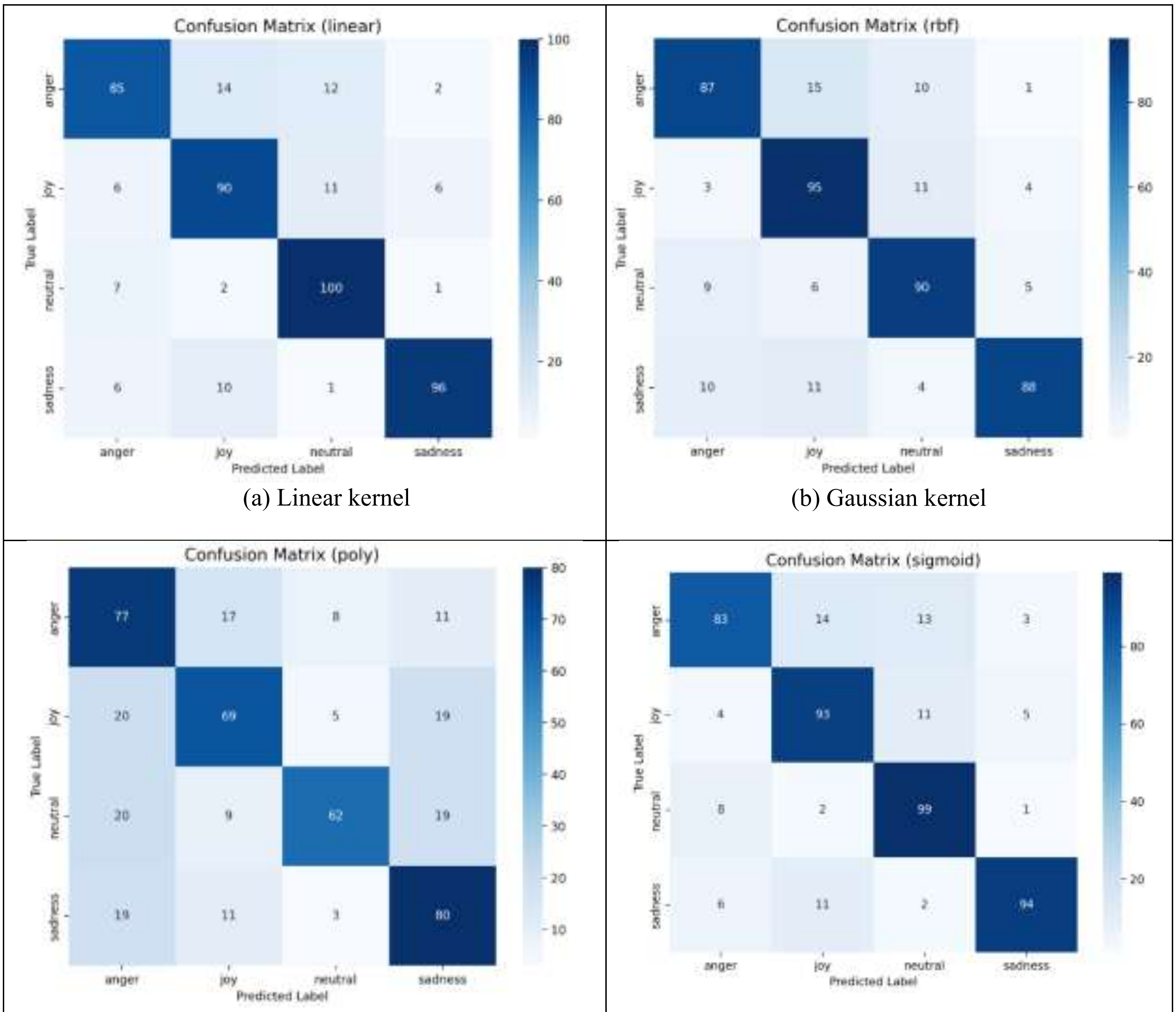
Table 5. Evaluation SVM Method

Kernel	Accuracy	Precision	Recall	F1-Score
Linear	0.8263	0.8286	0.8263	0.8260
Gaussian/RBF	0.8018	0.8069	0.8018	0.8024
Polynomial	0.6414	0.6571	0.6414	0.6422
Sigmoid	0.8218	0.8256	0.8218	0.8216

*Anik Vega Vitianingsih

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Issue 1, Januari 2026
DOI : https://doi.org/10.33395/sinkron.v10i1.15644
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The Figures 4(a) to (d) show the visualizations of evaluation results for each model kernel's performance, such as Linear, Gaussian, Polynomial, and Sigmoid kernel, respectively.



(a) Linear kernel

(b) Gaussian kernel

(c) Polynomial kernel

(d) Sigmoid kernel

Fig 4. Confusion Matrix Model Kernel's Performance

DISCUSSION

The result of study collected a total of 4001 user reviews related to FirstMedia's service concerns from the Google Play Store and the X app. After preprocessing and lexicon-based labeling, 3,069 reviews were retained with an imbalanced label distribution, consisting of 1,065 neutral, 748 anger, 692 joy, and 564 sadness reviews The findings show that the neutral category is the most dominant, indicating that users generally express their experiences or complaints in an informative manner without strong emotional intensity. After applying undersampling, a balanced dataset of 2,256 reviews (564 reviews per class) was obtained to reduce majority-class bias and support more stable multi-class classification.

The accuracy, precision, recall, and F1-score values of 0.8263, 0.8286, 0.8263, and 0.8260 respectively, demonstrate that the Linear kernel outperforms all other kernels. In comparison, the Gaussian, Polynomial, and Sigmoid kernels achieved accuracy scores of 0.8018, 0.6414, and 0.8218; precision scores of 0.8069, 0.6571, and 0.8256; recall scores of 0.8018, 0.6414, and 0.8218, and f1-score of 0.8024, 0.6422, 0.8216 respectively. These results indicate that the Linear kernel performs best when processing data with linear feature relationships, surpassing the Gaussian and Sigmoid kernels, which both outperform the Polynomial kernel. Overall, the Linear

*Anik Vega Vitianingsih

kernel provides the most reliable performance for this task, while the other kernels show less consistent and more ovariable results.

This result is consistent with the characteristics of TF–IDF features, which produce sparse and high-dimensional vectors where classes are often linearly separable. Under this representation, the Linear SVM is well suited because it can construct a robust decision boundary without requiring complex non-linear transformations. Furthermore, combining lexicon-based labeling with SVM can be effective because the lexicon provides structured emotion signals at the word level, while SVM learns discriminative patterns across reviews based on weighted term importance. However, because emotion labels were obtained through lexicon-based scoring, the model performance reflects agreement with proxy labels and may fail in cases such as sarcasm, contextual ambiguity, or domain-specific expressions not covered by the lexicon.

Conceptually, deep learning models (e.g., LSTM or Transformer-based approaches) could better capture contextual semantics, but they typically require larger human-annotated datasets and higher computational resources; therefore, the lexicon-based labeling combined with SVM offers a practical and scalable alternative for large-scale emotion analysis.

CONCLUSION

This study concludes that most user reviews fall into the neutral category, indicating that users generally describe their experiences in an informative and non-emotional manner. Following the neutral category, the most dominant emotions identified were anger, which commonly appeared in complaints regarding connection issues and slow customer response; sadness, which was associated with reviews mentioning declining service quality; and joy, which emerged in reviews praising network stability and satisfactory service.

The classification results further show that the SVM method with a Linear kernel is the most effective model for identifying user emotions, achieving the highest performance among all kernels with an accuracy of 82.63 %, a precision of 82.86 %, a recall of 82.63%, and an F1-score of 82.60 %. The model successfully classified reviews into four emotional categories joy, anger, sadness, and neutral providing a more comprehensive understanding of user perceptions toward service quality. Future study should focus on augmenting the dataset, implementing deep learning models, and employing more sophisticated balancing techniques to enhance emotion classification performance and more accurately reflect the subtleties of Indonesian-language literature.

However, this study has several limitations. First, emotion labels were derived from the NRC Indonesian Emotion Lexicon rather than manual human annotation, meaning that the model learned proxy labels and may be affected by lexicon bias and contextual ambiguity. Second, model evaluation was conducted using a single hold-out split, which may limit the robustness and generalizability of the reported performance. Future work should incorporate human-annotated emotion labels and repeated validation schemes (e.g., cross-validation) to improve reliability and reduce potential bias.

Future work should incorporate human-annotated emotion labels and repeated validation schemes (e.g., cross-validation) to improve reliability and reduce potential bias. In addition, hybrid approaches may be explored by integrating deep learning-based contextual embeddings with lexicon-based priors to better capture nuanced emotional expressions and enhance model robustness.

REFERENCES

Angelie Tania, V. E., & Oetama, R. S. (2025). Nusantara capital city sentiment analysis using support vector machine and logistic regression. *Indonesian Journal of Electrical Engineering and Computer Science*, *38*(3), 1708. https://doi.org/10.11591/ijeecs.v38.i3.pp1708-1721

Anggoro, D. A., & Permatasari, D. (2023). Performance comparison of the kernels of support vector machine algorithm for diabetes mellitus classification. *International Journal of Advanced Computer Science and Applications*, *14*(1), 580–585. https://doi.org/10.14569/IJACSA.2023.0140163

Berutu, S. S., Budiati, H., & Gulo, F. (2023). Data preprocessing approach for machine learning-based sentiment classification. *JURNAL INFOTEL*, *15*(4), 317–325. https://doi.org/10.20895/infotel.v15i4.1030

Chang, J., Chen, L., & Lin, L. (2021). A novel cluster based Over-sampling approach for classifying imbalanced sentiment data. *IAENG International Journal of Computer Science*, *48*(4), 1118–1128.

Damayanti, E., Vitianingsih, A. V., Kacung, S., & Cahyono, D. (2024). Sentiment analysis of Alfagift application user reviews using long short-term memory (LSTM) and support vector machine (SVM) Methods. *DECODE: JURNAL PENDIDIKAN TEKNOLOGI INFORMASI*, *4*(2), 509–521. https://doi.org/http://dx.doi.org/10.51454/decode.v4i2.478

Danyal, M. M., Khan, S. S., Khan, M., Ullah, S., Ghaffar, M. B., & Khan, W. (2024). Sentiment analysis of movie reviews based on NB approaches using TF–IDF and count vectorizer. *Social Network Analysis and Mining*, *14*(1). https://doi.org/10.1007/s13278-024-01250-9

Garc, R. A., Luna-garc, H., Celaya-padilla, J. M., Garc, A., Reveles-g, L. C., Flores-chaires, L. A., Delgado-contreras, J. R., Rondon, D., & Villalba-condori, K. O. (2024). A systematic literature review of modalities,

*Anik Vega Vitianingsih

e-ISSN : 2541-2019
p-ISSN : 2541-044X

trends, and limitations in emotion recognition, affective computing, and sentiment analysis. *Applied Sciences*, *14*(16). https://doi.org/10.3390/app14167165

Helmud, E., Helmud, E., Fitriyani, F., & Romadiana, P. (2024). Classification comparison performance of supervised machine learning random forest and decision tree algorithms using confusion matrix. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, *13*(1), 92–97. https://doi.org/10.32736/sisfokom.v13i1.1985

Hendrastuty, N., Isnain, A. R., Rahmadhani, A. Y., Studi, P., Informasi, S., Indonesia, U. T., Informatika, P. S., Indonesia, U. T., & Lampung, K. B. (2021). Analisis sentimen masyarakat terhadap program kartu prakerja pada twitter dengan metode support vector machine. *Jurnal Informatika: Jurnal Pengembangan IT (JPIT)*, *6*(3), 150–155. https://doi.org/10.30591/jpit.v6i3.2870

Iswaratama, A. (2024). Peran komunitas virtual dalam mendorong interaksi sosial di era digital. *HISTORICAL: Journal of History and Social Sciences*, *3*(1), 51–61. https://doi.org/10.58355/historical.v3i1.100 Vol.

Jaya, S. I. A., & Zeniarja, J. (2024). Sentiment analysis of genshin impact on X: Mental health implications using TF-IDF and support vector machine. *Sinkron*, *8*(3), 1589–1599. https://doi.org/10.33395/sinkron.v8i3.13716

March Vircan Karuna, H. M. (2023). Analisis sentiment review kepuasan pengguna wi-fi First Media di Twitter. *Journal of Management and Bussines (JOMB)*, *183*(2), 153–164. https://doi.org/10.31539/jomb.v5i2.8301 ANALISIS

Palomino, M. A., & Aider, F. (2022). Evaluating the effectiveness of text pre-processing in sentiment analysis. *Applied Sciences*, *12*, 8765. https://doi.org/10.3390/app12178765

Praghakusma, A. Z., & Charibaldi, N. (2021). Komparasi fungsi kernel metode Support Vector Machine untuk Analisis Sentimen Instagram dan Twitter (Studi Kasus : Komisi Pemberantasan Korupsi). *JSTIF (Jurnal Sarjana Teknik Informatika)*, *9*(2), 88. https://doi.org/10.12928/jstie.v9i2.20181

Prasetio, M. D., Xavier, R. Y., Rachmat, H., Wiyono, W., & Atmaja, D. S. E. (2021). Sentiment analysis on myindihome user reviews using support vector machine and naïve bayes classifier method. *International Journal of Industrial Optimization*, *2*(2), 141. https://doi.org/10.12928/ijio.v2i2.4449

Putra Selian, R. I., Vitianingsih, A. V., Kacung, S., Lidya Maukar, A., & Febrian Rusdi, J. (2024). Sentiment analysis of public responses on social media to satire joke using naive bayes and KNN. *Sinkron*, *8*(3), 1443–1451. https://doi.org/10.33395/sinkron.v8i3.13721

Qi, Y., & Shabrina, Z. (2023). Sentiment analysis using Twitter data: A comparative application of lexicon- and machine-learning-based approach. *Social Network Analysis and Mining*, *13*(1), 1–14. https://doi.org/10.1007/s13278-023-01030-x

Rahmadani, R., Rahim, A., & Rudiman, R. (2024). Analisis sentimen ulasan "Ojol the Game" di Google Play Store menggunakan algoritma Naive Bayes dan model ekstraksi fitur TF-IDF untuk meningkatkan kualitas game. *Jurnal Informatika Dan Teknik Elektro Terapan*, *12*(3). https://doi.org/10.23960/jitet.v12i3.4988

Shamsi, M., & Beheshti, S. (2025). Separability and scatteredness (S&S) ratio-based efficient SVM regularization parameter, kernel, and kernel parameter selection. *Pattern Analysis and Applications*, *28*(1), 1–21. https://doi.org/10.1007/s10044-025-01411-2

Simay Akar, S. A., Yang Sok Kim, Y. S. K., & Mi Jin Noh, M. J. N. (2024). Sentiment analysis on "HelloTalk" app reviews using NRC Emotion Lexicon and GoEmotions dataset. *Korean Institute of Smart Media*, *13*(6), 35–43. https://doi.org/10.30693/smj.2024.13.6.35

Surya, A., Ansyah, S., Kurniawan, A. P., Kholifah, A. N., & Purwitasari, D. (2023). A hybrid method on emotion detection for Indonesian tweets of COVID-19. *JURNAL RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *7*(2), 254–262. https://doi.org/10.29207/resti.v7i2.4816

Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, *13*(7). https://doi.org/10.3390/app13074550

Vanacore, A., Pellegrino, M. S., & Ciardiello, A. (2024). Fair evaluation of classifier predictive performance based on binary confusion matrix. *Computational Statistics*, *39*(1), 363–383. https://doi.org/10.1007/s00180-022-01301-9

Venugopal, J. P., Subramanian, A. A. V., Sundaram, G., Rivera, M., & Wheeler, P. (2024). A comprehensive approach to bias mitigation for sentiment analysis of social media data. *Applied Sciences*, *14*(23), 1–32. https://doi.org/10.3390/app142311471

Zad, S., Jimenez, J., & Finlayson, M. A. (2021). Hell hath no fury? correcting bias in the NRC Emotion Lexicon. *International Committee on Computational Linguistics (ICCL)*, 102–113. https://doi.org/10.18653/v1/2021.woah-1.11

*Anik Vega Vitianingsih