SENTIMENT ANALYSIS OF DIGITAL POPULATION IDENTITY APPLICATION: ANALYSIS OF NAÏVE BAYES AND SUPPORT VECTOR MACHINE

Fajar Alpinnanto[1]; Mardiana Purwaningsih[2*]; Alrafiful Rahman[3]

Faculty of Information Technology[1,2,3]
Perbanas Institute, Jakarta, Indonesia[1,2,3]
https://perbanas.id/[1,2,3]
fajar.alpinnanto09@perbanas.id[1], mardiana@perbanas[2*], alrafiful.rahman@perbanas.id[3]

(*) Corresponding Author

The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.

Abstract— *Digital Population Identity (Identitas Kependudukan Digital/IKD) is a digital identity system developed by the Indonesian government to facilitate citizens' access to public services through digital platforms. This study compares the performance of two machine learning algorithms, Naïve Bayes and Support Vector Machine (SVM), in classifying user reviews of the IKD application collected from the Google Play Store. The review data were preprocessed through text cleaning, case folding, tokenization, stopword removal, and stemming to improve text quality before classification. To address the issue of class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied to enhance the performance of the classification models. The experimental results show that both algorithms are effective in classifying positive and negative sentiments; however, the SVM model outperformed Naïve Bayes in terms of classification performance. Furthermore, the application of SMOTE contributed to improving the accuracy of sentiment classification by balancing the distribution of the training data. These findings provide valuable insights for IKD application developers to better understand user feedback and support continuous improvements in application quality and user experience.*

Keywords: *Digital Population Identity, E-Government, Naive Bayes, Sentiment Analysis, Support Vector Machine.*

Intisari— Identitas Kependudukan Digital (IKD) adalah sistem identitas digital yang dikembangkan oleh pemerintah Indonesia untuk memfasilitasi akses warga negara terhadap layanan publik melalui platform digital. Studi ini membandingkan kinerja dua algoritma pembelajaran mesin, Naïve Bayes dan Support Vector Machine (SVM), dalam mengklasifikasikan ulasan pengguna aplikasi IKD yang dikumpulkan dari Google Play Store. Data ulasan diproses terlebih dahulu melalui pembersihan teks, case folding, tokenisasi, penghapusan stopword, dan stemming untuk meningkatkan kualitas teks sebelum klasifikasi. Untuk mengatasi masalah ketidakseimbangan kelas, Synthetic Minority Over-sampling Technique (SMOTE) diterapkan untuk meningkatkan kinerja model klasifikasi. Hasil eksperimen menunjukkan bahwa kedua algoritma efektif dalam mengklasifikasikan sentimen positif dan negatif; namun, model SVM mengungguli Naïve Bayes dalam hal kinerja klasifikasi. Lebih lanjut, penerapan SMOTE berkontribusi pada peningkatan akurasi klasifikasi sentimen dengan menyeimbangkan distribusi data pelatihan. Temuan ini memberikan wawasan berharga bagi pengembang aplikasi IKD untuk lebih memahami umpan balik pengguna dan mendukung peningkatan berkelanjutan dalam kualitas aplikasi dan pengalaman pengguna.

Kata Kunci: *Identitas Kependudukan Digital, Pemerintahan Elektronik, Naive Bayes, Analisis Sentimen, Support Vector Machine.*

INTRODUCTION

Various components of public services have transitioned to digital platforms as a result of advancements in information technology, with the population management system being a prime example. One prominent innovation is the implementation of the Digital Population Identity (IKD) application, which enables the public to access civil registration services and information online, swiftly, and efficiently. This application not only simplifies administrative processes but also serves as a two-way communication medium

between the government and the community. Driven by the high adoption rate of digital technology among the public, various public opinions and responses regarding these services have emerged, captured through social media platforms and application reviews. These data serve as an essential resource for evaluating service quality based on user perception (Widodo & Kusnan, 2023).

In this context, sentiment analysis becomes a highly relevant approach to systematically and quantitatively interpret user opinions. Through sentiment analysis, emotional patterns and opinion tendencies—whether positive, negative, or neutral—can be identified, allowing developers and policymakers to formulate more targeted service improvement strategies (Nur Afrilia et al., 2024). However, to obtain accurate results, reliable text classification algorithms are required. Naive Bayes (NB) and Support Vector Machine (SVM) are two of the most commonly utilized sentiment classification techniques. Naive Bayes is renowned for its simplicity and efficiency in processing text data (Rahayu et al., 2022), while SVM is recognized for its strong generalization capability and high accuracy across various application domains (Kurniawan et al., 2023).

Although both algorithms have been widely used in prior research, their effectiveness highly depends on data characteristics and the implementation context. This study focuses on comparing the performance of Naive Bayes and SVM in classifying the sentiment of IKD application user reviews. This comparison is significant given the high linguistic variation in digital reviews, including the use of informal language and non-standard terms that frequently appear in user opinions (Soraya & Indriyanti, 2024).

To address class imbalance, a data balancing approach was employed to reduce the dominance of a particular class and improve the learning capability of the classification models. This imbalance can hinder the performance of classification models and lead to a bias toward the majority class (Rina Noviana & Isram Rasal, 2023). To address this issue and balance the class distribution, the Synthetic Minority Over-sampling Technique (SMOTE) is employed (Siringoringo, 2018). This technique has been proven to enhance model accuracy and improve classification sensitivity (Sulistiyowati & Jajuli, 2020).

Based on this review, it can be concluded that the utilization of classification algorithms such as Naive Bayes and Support Vector Machine (SVM), alongside class imbalance handling techniques like SMOTE, is crucial for improving the accuracy of sentiment analysis, particularly on digital public service application review data. This study aims to evaluate and compare the performance of Naive Bayes and SVM algorithms in analyzing the sentiment of Digital Population Identity (IKD) application reviews, as well as to examine the impact of applying SMOTE to mitigate class imbalance issues. The evaluation is conducted by considering performance metrics such as accuracy, precision, recall, and F1-score. The results of this study are expected to provide theoretical and practical contributions to the development of text classification systems within the digital public service domain utilizing machine learning. Furthermore, these findings are anticipated to assist application developers and policymakers in making informed decisions to enhance services based on public opinion.

MATERIALS AND METHODS

A. Literature Review

This stage encompasses the collection of peer-reviewed articles and academic journals relevant to the current study. The literature focus includes sentiment analysis of user reviews for the Digital Population Identity (IKD) application utilizing Naive Bayes and Support Vector Machine, alongside other studies within identical domains.

Prior research has extensively investigated the implementation of machine learning algorithms for sentiment analysis. For instance, a previous study employed the Support Vector Machine (SVM) algorithm to analyze public perceptions regarding the Digital Population Identity (IKD) application on Facebook. The study revealed that the majority of user reviews were negative, with the SVM classification model achieving an accuracy of 77% (Lestari et al., 2023). Conversely, another study compared both algorithms utilizing Twitter data related to the University of Muhammadiyah Malang; however, their findings indicated that Naive Bayes outperformed SVM with an accuracy of 73.65% compared to 71.25% (Fikri et al., 2020). In contrast, similar empirical evidence confirmed the superiority of SVM within the context of the Pluang digital financial application, yielding an accuracy of 99.50%, which slightly surpassed Naive Bayes at 99.25% (Maulana et al., 2024).

Furthermore, several literatures have highlighted the critical importance of mitigating data imbalance issues. One study integrated Naive Bayes with the Synthetic Minority Over-sampling Technique (SMOTE) to enhance class distribution within credit datasets (Sulistiyowati & Jajuli, 2020), while another study demonstrated that combining SMOTE with the Nearest Neighbor algorithm significantly improves classification accuracy on imbalanced data (Siringoringo, 2018). Additionally, research evaluated user sentiments toward the

PeduliLindungi application using a Naive Bayes Classifier, revealing that the majority of the reviews were negative, with a recorded accuracy of 83.3% (Locarso, 2022). In a similar vein, research focused on the sentiment analysis of various online shopping applications on the Google Play Store using SVM indicated that accuracy fluctuates depending on data characteristics, reaching its peak for Tokopedia at 90.67% (Ahmadi et al., 2020). Complementing these classification methodologies, another study implemented the K-Means clustering technique to analyze user activity patterns on social media, which provides valuable insights into grouping digital behavioral patterns (Nur Afrilia et al., 2024).

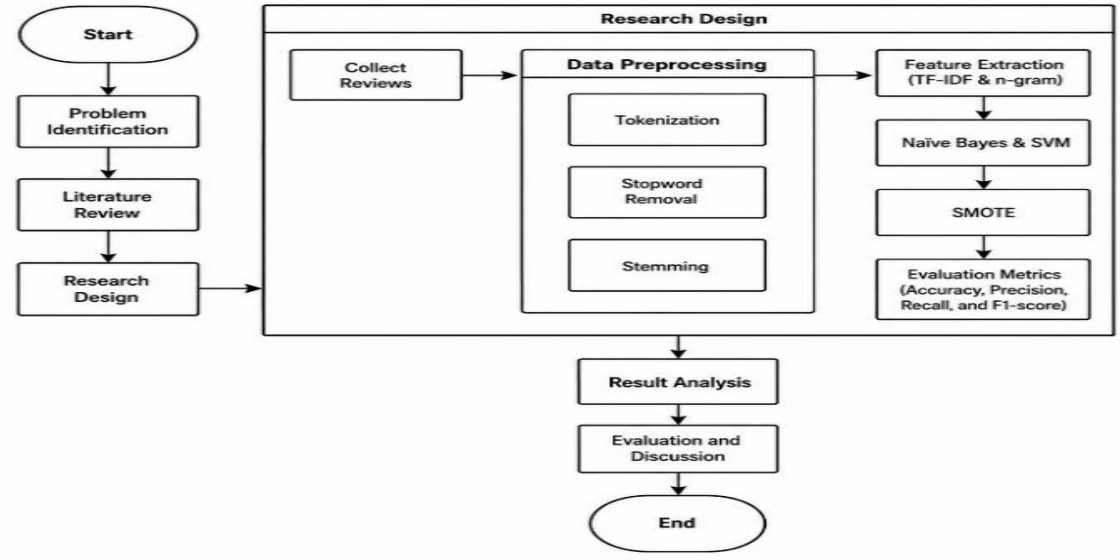
B. Problem Definition and Data Acquisition

Identifying the research subject constitutes the foundational step of this study. Research problems typically stem from various real-world case studies, with the digital public service domain serving as a prime example. In this study, two supervised learning algorithms are utilized to analyze user sentiment toward the Digital Population Identity (IKD) application. Naive Bayes and Support Vector Machine (SVM) were selected due to their simplicity and computational efficiency in text classification, with SVM specifically chosen for its robust capability to handle complex and high-dimensional text data. To construct the dataset, the Python library google_To construct the dataset, the Python library google_play_scraper was employed to collect user reviews of the Digital Population Identity (IKD) application from the Google Play Store. A total of 10,000 reviews posted between January and February 2026 were initially retrieved. The collected reviews then underwent several preprocessing stages, including data cleaning, case folding, tokenization, stopword removal, stemming, and the removal of duplicate, empty, and irrelevant records. As a result of these preprocessing steps, only valid reviews were retained and used as the final dataset for sentiment classification and model evaluation.

C. Research Framework

This study employs a quantitative approach through experimental research as a structured procedure to collect, analyze, and compare the effectiveness of two text classification algorithms Naive Bayes and Support Vector Machine (SVM) in analyzing user review sentiments of the Digital Population Identity (IKD) application. To mitigate the class imbalance problem within the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) is applied, allowing the models to learn optimally from both the positive and negative classes (M. S. Rahman et al., 2023). A systematic comparison between the two algorithms is conducted to yield a valid analysis, considering that both possess distinct data processing characteristics and require appropriate handling when applied to real-world review data. Fig. 1 illustrates the sequential phases representing the overall research workflow.

At this stage, the research framework is established to outline the systematic workflow of the study, utilizing a quantitative approach to compare the performance metrics of both algorithms. The detailed research workflow illustrated in Fig. 1 is described as follows (Vincent et al., 2024):



Source: (Research Results, 2026)

Figure 1. Research Methodology Flowchart

1. Data Acquisition: The required dataset for this study is gathered during this initial phase. The collected data consists of user reviews of the Digital Population Identity (IKD) application extracted from the Google Play Store platform, which are categorized into neutral, positive, and negative sentiments.
2. Data Preprocessing: This phase involves comprehensive data cleaning, which eliminates redundant characters, including numbers and punctuation marks. It is followed by tokenization to segment the raw text into individual words or tokens. Subsequently, stopword removal is executed to filter out common words that carry minimal semantic value in sentiment analysis. Stemming is then applied to reduce inflected or derived words into their respective base forms. Finally, feature extraction is conducted using TF-IDF and n-grams to capture relevant features from the text corpus.
3. Model Development: This stage involves model training, where both the Naïve Bayes and Support Vector Machine (SVM) algorithms are trained using the preprocessed dataset to learn the underlying patterns of user sentiment. To address the issue of class imbalance, a data balancing process is incorporated into the training stage. Specifically, the Synthetic Minority Over-sampling Technique (SMOTE) is applied to generate synthetic samples for the minority class, resulting in a more balanced distribution of training instances. This approach aims to reduce the bias caused by the dominance of one class and improve the models' ability to learn representative patterns across all classes, thereby enhancing the reliability of the subsequent sentiment classification results.
4. Model Evaluation: The trained models are validated using a testing dataset to measure their classification effectiveness. Performance evaluation metrics include the F1-score, accuracy, precision, and recall.

D. Performance Analysis

The analysis of the results focuses on an algorithm comparison by evaluating the accuracy, precision, recall, and F1-score yielded by both Naive Bayes and SVM. Furthermore, the impact of SMOTE is assessed to determine the extent to which it enhances model performance, particularly in mitigating class imbalance issues. The analysis also evaluates the influence of TF-IDF and n-gram feature extraction techniques on the final sentiment classification outcomes. The mathematical formulations used to compute accuracy, precision, recall, and F1-score are defined as follows (A. Rahman & Bustamam, 2022):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (4)$$

Where:
TP (True Positive): The number of correctly predicted positive instances.
TN (True Negative): The number of correctly predicted negative instances.
FP (False Positive): The number of incorrectly predicted instances, where negative samples are classified as positive.
FN (False Negative): The number of incorrectly predicted instances, where positive samples are classified as negative.

E. Evaluation and Discussion

This phase assesses the final outputs of the classification models based on the selected performance metrics. It involves a comprehensive comparison of the obtained results, an in-depth discussion on the respective strengths and limitations of each algorithm, and an analysis of the specific impact of SMOTE in enhancing the overall performance of the models.

RESULTS AND DISCUSSION

This study aims to evaluate and compare the performance of two text classification algorithms Naive Bayes and Support Vector Machine (SVM) in conducting sentiment analysis on user reviews of the Digital Population Identity (IKD) application. The analytical pipeline encompasses data acquisition, preprocessing, labeling, feature extraction, and model evaluation, implemented both with and without the integration of the Synthetic Minority Over-sampling Technique (SMOTE).

A. Data Labeling

Following the analysis of the sentiment distribution, the fully processed and labeled dataset was exported and archived in a CSV format to facilitate subsequent modeling phases, including feature extraction, model training, and performance

evaluation. This file was designated as ulasan_ikd_labeled.csv and saved using the UTF-8-SIG encoding scheme to ensure cross-platform compatibility when accessed via spreadsheet software such as Microsoft Excel. Archiving this labeled dataset provides a structured foundation to proceed to the subsequent stages of the sentiment analysis workflow, specifically the training of machine learning models to automatically classify sentiments based on user review texts.

Table 1. Dataset Samples in Indonesian After the Labeling Process

Username	Rating Score	Raw Review	Preprocessed Review	Sentiment Label
YS	5	bagus banget 👍 membantu 👍	bagus bangetmembantu	Positive
Mm	1	ancur parah sistemnya, masa verifikasi wajah susah seteng mati	ancur parah sistem masa verifikasi wajah susah teng mati	Negative
HS	5	Alkhamdulillah, terima kasih 🙏	alkhamdulillah terima kasih	Positive
Fa	1	gaguna udah daftar di suruh daftar lagi	gaguna udah daftar suruh daftar	Negative
GP	2	Kalo Udah Lupa PIN yaudah ilang deh gak bisa buka ktp di aplikasi ini untuk selamanya 😅	kalo udah lupa pin yaudah ilang deh gak buka ktp aplikasi lama	Negative

Source: (Research Results, 2026)

In this study, the data labeling process was executed automatically based on the user rating scores. To establish a binary classification framework, reviews with ratings of 4 and 5 were categorized as positive, whereas those with ratings of 1 and 2 were classified as negative. Neutral reviews, represented by a rating score of 3, were deliberately excluded from the dataset. Consequently, the fully classified dataset comprises two distinct classes with the following distribution: Negative: 6,779 instances and Positive: 2,745 instances

B. Feature Extraction

Feature extraction was executed by converting the textual data into high-dimensional, sparse numerical matrices. This sparsity implies that the majority of the matrix elements contain zero values, as each individual document only incorporates a small fraction of the entire vocabulary. This characteristic is widely inherent in text representation frameworks and is particularly efficient for processing by sparse-matrix-based models such as Naive Bayes and SVM. The following table summarizes the outputs of the TF-IDF feature extraction process across both configurations (Table 2).

Table 2. Feature Extraction Results

Extraction Type	Total Instances	Number of Features	Density (%)	Matrix Type
TF-IDF Unigram	9292	3000	0,292%	scipy.sparse._csr.csr_matrix
TF-IDF Bigram	9292	3000	0,367%	scipy.sparse._csr.csr_matrix

Source: (Research Results, 2026)

The extraction results demonstrate that both text representation methods yield highly sparse matrices, indicating that only a minor fraction of the total features are present within each individual document. Nevertheless, these representations remain highly effective as they capture critical structural patterns and semantic contexts within the text, while simultaneously preserving computational efficiency during the training and evaluation phases of the sentiment classification models.

C. Model Implementation: Naive Bayes and SVM

The evaluation results showed that the SVM model with the bigram configuration performed best among all tested combinations (Table 3). This model achieved an accuracy of 90.26%, a precision of 90.13%, a recall of 90.26%, and an F1-score of 90.15%. This means that the SVM was not only able to consistently recognize sentiment patterns in the data, but also effectively balanced the ability to detect both positive and negative classes fairly.

In more detail, the SVM (bigram) demonstrated excellent ability to classify the negative class with an F1-score of 93.26%, and was quite solid in recognizing the positive class with an F1-score of 82.44%. This indicates that the model is highly effective in recognizing negative reviews, which are generally more common in public service application datasets, while also being quite responsive to the less frequent positive reviews.

Conversely, Naïve Bayes, in both the unigram and bigram configurations, tended to have high

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 2 September 2026 | DOI https://doi.org/10.33480/techno.v23i2.8520

recall for the negative class but were weak in recognizing the positive class. In the unigram configuration, Naïve Bayes recorded an accuracy of 87.90% with an F1-score of 87.11%. Meanwhile, in the bigram configuration, its performance decreased slightly with an accuracy of 87.68% and an F1-score of 87.01%. This can be explained by the basic nature of the Naïve Bayes algorithm which assumes independence between features, which in reality is not always the case in complex natural language contexts such as user reviews.

In conclusion, these experiments demonstrate that the selection of classification algorithms and feature extraction techniques substantially influences sentiment classification outcomes. The combination of SVM and TF-IDF bigrams proved to be the most effective approach and is highly recommended for identical research domains, specifically for sentiment analysis of user reviews within public service digital platforms such as IKD.

Table 3. Evaluation Results of Naive Bayes and SVM Models

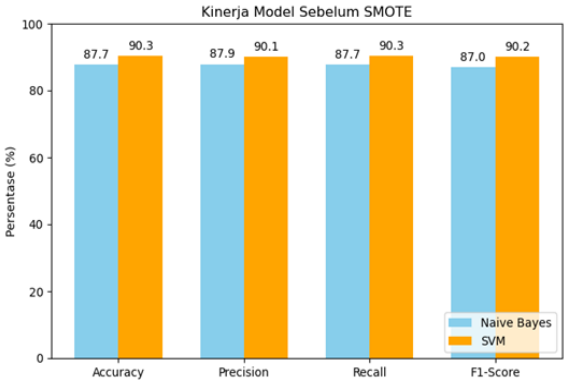
Model	Accuracy	Precision	Recall	F1-Score
Naïve Bayes (Unigram)	87,90 %	88,43%	87,90%	87,11 %
SVM (Unigram)	89,30 %	89,18%	89,30%	89%
Naïve Bayes (Bigram)	87,68 %	87,88%	87,58%	87,01 %
SVM (Unigram)	90,26 %	90,13%	90,26%	90,15

Source: (Research Results, 2026)

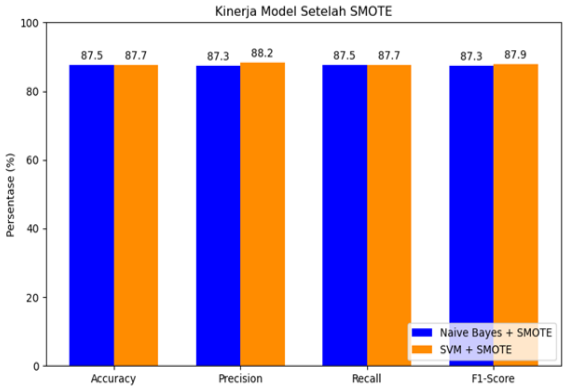
D. Implementation of the SMOTE Technique

Class imbalance within the review distribution represents a primary challenge in sentiment analysis. In this study, the negative class contains a significantly larger volume of instances compared to the positive class. This skewed distribution typically causes machine learning models to become biased toward accurately recognizing the majority class (negative), while yielding poorer performance on the minority class (positive). To mitigate this issue, the Synthetic Minority Over-sampling Technique (SMOTE) was implemented to balance the data distribution across classes.

Before applying the data balancing technique, the initial classification models were trained and evaluated using the original dataset, which exhibited an imbalanced class distribution. Under this condition, both the Naïve Bayes and SVM models tended to be influenced by the dominant class, potentially limiting their ability to learn minority-class patterns effectively. Figure 3 presents the baseline performance of both models before the balancing process, using accuracy, precision, recall, and F1-score as evaluation metrics. The SVM model achieved higher performance than Naïve Bayes across all metrics, with an accuracy of 90.3% and an F1-score of 90.2%, while Naïve Bayes obtained an accuracy of 87.7% and an F1-score of 87.0%. Although these baseline results indicate good overall classification performance, they do not necessarily reflect balanced predictive capability across all classes due to the unequal class distribution.



Source: (Research Results, 2026)
Figure 2. Performance Comparison Before SMOTE



Source: (Research Results, 2026)
Figure 3. Performance Comparison After SMOTE

After applying the SMOTE technique to balance the class distribution in the training dataset, both classification models were retrained using the balanced data. The evaluation results on the testing dataset are presented in Fig. 3, illustrating the models' performance after the data balancing process. Compared with the baseline results, both models demonstrated more balanced classification performance, indicating a reduced influence of class dominance during the learning process. Following the application of SMOTE, Naïve Bayes achieved an accuracy of 87.5% and an F1-score of 87.3%, while SVM obtained an accuracy of 87.7% and an F1-score of 87.9%. Although the overall accuracy and F1-

score of the SVM model decreased slightly compared with the pre-balancing results, the model exhibited a more balanced classification capability across all classes, suggesting that the learning process became less biased toward the previously dominant class.

In general, the integration of SMOTE exerts a positive impact on inter-class classification balance, particularly within the framework of imbalanced datasets. SMOTE successfully improves the capability of Naive Bayes to identify positive sentiments without suffering a severe degradation in performance on the negative class. Meanwhile, despite experiencing a slight reduction in aggregate scores, the SVM model demonstrated greater consistency across a more equitable distribution. Therefore, within the context of digital public services such as the IKD application, the deployment of SMOTE is highly recommended to enhance the representation of minority sentiment trends in user feedback.

E. Analysis of Results

This section evaluates the performance outputs of the two classification models deployed in this study: Naive Bayes and SVM. Through these models, user perspectives extracted from the Digital Population Identity (IKD) application reviews are rigorously analyzed. Furthermore, this analysis examines the efficacy of the Synthetic Minority Over-sampling Technique (SMOTE) in enhancing model performance to mitigate class imbalance within the dataset. To ensure a comprehensive evaluation, three primary metrics are analyzed: precision, recall, and the F1-score. The results demonstrate that the integration of SMOTE successfully resolves the class imbalance issue and substantially improves the models' capability to classify sentiments accurately, particularly for the minority class, which was previously prone to misclassification.

Following the implementation of both models, the next phase involves comparing their performance based on precision, recall, and F1-score. The evaluation yields a comprehensive insight into the respective strengths and limitations of each model in handling positive and negative sentiments.

1) Precision

Precision measures the extent to which a model correctly classifies positive instances, indicating the proportion of reviews predicted as a specific class that genuinely belong to that class. In this study, Naive Bayes demonstrated a higher precision on the negative class, indicating its efficacy in accurately identifying reviews containing negative sentiment. However, following the application of SMOTE, the Support Vector Machine (SVM) exhibited superior precision for the positive class. In conclusion, SVM is more capable of handling class imbalance effectively, particularly after the minority class dataset is expanded using SMOTE.

The superiority of Naive Bayes in terms of precision for the negative class can be attributed to its fundamental assumption of feature independence. This assumption performs well on text data containing distinct and explicit keywords for a specific class. Conversely, for the minority class, which often exhibits higher ambiguity, Naive Bayes appears to be less effective than SVM, as the latter is capable of capturing a broader context within the textual data.

2) Recall

Recall quantifies the model's capability to capture all relevant instances within each respective class. In this context, the recall for the minority class experienced a substantial increase after the implementation of SMOTE across both models. SVM exhibited a more pronounced surge in recall for the minority class post-SMOTE, demonstrating that SVM is more effective in classifying minority sentiment reviews once the dataset is balanced. The deployment of SMOTE generates synthetic data that resolves class imbalance; since SVM inherently manages complex data distributions efficiently, the model successfully enhances the recall of the minority class.

In contrast, Naive Bayes yielded a lower recall for the minority class despite the integration of SMOTE. This limitation stems from the conditional independence assumption of Naive Bayes, which restricts its capacity to handle inter-feature dependencies that are crucial for capturing complex and ambiguous patterns within minority sentiment reviews.

3) F1-Score

The F1-score serves as an equitable metric that provides a comprehensive overview of model performance by harmonizing precision and recall for each class. Higher F1-scores achieved by SVM across both negative and positive classes underscore its robust performance in mitigating class imbalance and generating well-balanced predictions.

The F1-score comparison demonstrates that while Naive Bayes is effective for the majority negative class, SVM offers a superior capacity to handle the minority positive class, particularly after SMOTE implementation (Table 4). This indicates that the Support Vector

Machine is more resilient against class imbalance and more proficient in learning complex patterns within textual datasets.

Table 4. Performance Evaluation Comparison Between Naive Bayes and SVM Models

Metric	Naive Bayes	Support Vector Machine
Precision	87,88%	90,13%
Recall	87,58%	90,26%
F1-Score	87,01%	90,15

Source: (Research Results, 2026)

SVM demonstrates superior performance in terms of precision, recall, and F1-score for the minority positive class, although Naive Bayes maintains a robust performance for the negative class.

F. Discussion

This section discusses the operational mechanics of both models within the context of the deployed dataset and evaluates how the implementation of the SMOTE technique mitigates the class imbalance problem inherent in the data. Furthermore, the obtained results are benchmarked against relevant published literature to provide a broader overview regarding the efficacy and applicability of the utilized models.

Naive Bayes is a widely adopted probabilistic model in text classification, particularly for datasets characterized by independent features. The experimental results indicate that the Naive Bayes model performs robustly on the negative class, where it effectively identifies highly indicative keywords associated with negative sentiment. Additionally, the Naive Bayes method proved to be more computationally efficient and faster than SVM during the training phase. This finding aligns with a study conducted by (NURHASANAH & NAUFAL, 2024), which argues that Naive Bayes offers superior efficiency for large yet less complex datasets. In terms of implementation simplicity, Naive Bayes is straightforward to deploy and does not require extensive hyperparameter tuning, making it highly practical and fast. Consequently, Naive Bayes is highly suitable for applications requiring real-time classification, such as mobile-based platforms.

However, a primary limitation of Naive Bayes discussed herein pertains to its fundamental feature independence assumption. Although highly effective for the negative class, this method struggles to handle the [Opsi A: minority positive / Opsi B: neutral] class, which exhibits strong inter-word dependencies. For instance, in sentences with ambiguous sentiments, Naive Bayes fails to capture more complex contexts. This underscores the inefficiency of Naive Bayes when dealing with highly ambiguous data (Tarwoto et al., 2025). Furthermore, Naive Bayes demonstrates lower sensitivity toward class imbalance adjustments even after SMOTE is applied. The model consistently encounters difficulties in processing the [Opsi A: minority / Opsi B: neutral] class, which was underrepresented in the initial dataset. This indicates that Naive Bayes is more susceptible to class imbalance issues compared to other models like SVM.

Conversely, SVM is a more sophisticated algorithm that is highly effective for high-dimensional data, such as text analysis. Based on the experimental outcomes, SVM demonstrated superior capabilities in classifying the [Opsi A: minority / Opsi B: neutral] class, which was previously challenging for Naive Bayes to identify. Additionally, SVM possesses the capacity to handle non-linear data; in this study, it exhibited exceptional performance in managing highly complex datasets and diverse feature spaces. By leveraging the margin maximization concept to separate data, SVM manages inter-feature dependencies more effectively. This is consistent with the findings of (Siti Aisah et al., 2024), which state that SVM offers superior capabilities in resolving text data classification challenges.

SVM also demonstrated robust performance on the minority class. The integration of SMOTE yielded a significant increase in the recall and F1-score of the [Opsi A: minority / Opsi B: neutral] class. This indicates that SVM is more proficient at leveraging synthetic data generated by SMOTE to enhance model performance on underrepresented classes. SVM successfully captured [Opsi A: minority / Opsi B: neutral] sentiments more accurately than Naive Bayes post-SMOTE implementation, proving its effectiveness in resolving class imbalance. On the other hand, certain limitations of SVM must be anticipated, particularly regarding its training time complexity.

As noted in (Akrom, 2024), SVM requires substantial training time, especially when dealing with large and complex datasets. This was evident in this study, where SVM consumed more time and computational resources compared to Naive Bayes, which may pose a constraint for applications with strict training time limitations. Furthermore, SVM is highly sensitive to hyperparameter selection, such as C and γ, which must be meticulously tuned to achieve optimal results. Without proper optimization, SVM may yield suboptimal performance, as corroborated by (Yulian & Susanto, 2025), which emphasized that parameter tuning is critical for enhancing SVM model accuracy.

G. Research Implications

Based on the empirical findings of this study, it can be concluded that the Support Vector Machine (SVM) demonstrates superior performance in terms of precision, recall, and F1-score, particularly for the [Opsi A: minority / Opsi B: neutral] class following the application of SMOTE. The integration of SMOTE successfully enhances the inter-class balance within the dataset, which subsequently enables both models to identify [Opsi A: minority / Opsi B: neutral] sentiments with greater accuracy. Although Naive Bayes exhibits faster training times and superior computational efficiency, the Support Vector Machine proves to be more resilient against class imbalance and more effective in handling data characterized by high inter-feature dependencies. Consequently, for applications requiring high precision in analyzing complex sentiment patterns, the deployment of the Support Vector Machine is highly recommended. Conversely, Naive Bayes remains a viable alternative for applications prioritizing high processing speed and resource efficiency, albeit with a marginal trade-off in classification accuracy.

CONCLUSION

This study compared the performance of Naïve Bayes and Support Vector Machine (SVM) for sentiment classification of Digital Population Identity (IKD) application reviews using TF-IDF bigram features under both the original and balanced data distributions. The experimental results indicate that SVM consistently achieved slightly better overall performance than Naïve Bayes across the evaluation scenarios. The application of a data balancing technique contributed to reducing the effect of class imbalance, resulting in a more balanced classification performance across classes, although the overall performance metrics changed only marginally. These findings suggest that combining TF-IDF bigram features with an appropriate data balancing strategy can support more reliable sentiment classification of IKD application reviews. Future research may investigate alternative balancing techniques, feature extraction methods, and advanced machine learning or deep learning models using more diverse datasets to further validate these findings.

REFERENCE

Ahmadi, M. I., Apriani, F., Kurniasari, M., Handayani, S., & Gustian, D. (2020). Sentiment Analysis Online Shop on the Play Store Using Method Support Vector Machine (Svm. *Seminar Nasional …*, *2020*(Semnasif), 196–203. http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4101

Akrom, M. (2024). Komparasi Svm Klasik Dan Kuantum Dalam Klasifikasi Biner Biji Gandum (Seeds). *Networking Engineering Research Operation*, *9*(1), 49–58. https://doi.org/10.21107/nero.v9i1.28082

Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Comparison of Naïve Bayes and Support Vector Machine Methods in Twitter Sentiment Analysis. *Smatika Jurnal*, *10*(02), 71–76.

Kurniawan, I., Lia Hananto, A., Shofia Hilabi, S., Hananto, A., Priyatna, B., & Yuniar Rahman, A. (2023). Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter. *Jurnal Teknik Informatika Dan Sistem Informasi*, *10*(1), 731–740. http://jurnal.mdp.ac.id

Lestari, R., Erfina, A., & Jatmiko, W. (2023). Penerapan Algoritma Support Vector Machine pada Analisis Sentimen Terhadap Identitas Kependudukan Digital. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *10*, 1063–1070. https://doi.org/10.25126/jtiik.20231057264

Locarso, G. K. (2022). Analisis Sentimen Review Aplikasi Pedulilindungi Pada Google Play Store Menggunakan Nbc. *JTIK (Jurnal Teknik Informatika Kaputama)*, *6*(2), 353–361. https://doi.org/10.59697/jtik.v6i2.207

Maulana, B. A., Fahmi, M. J., Imran, A. M., & Hidayati, N. (2024). Analisis Sentimen Terhadap Aplikasi Pluang Menggunakan Algoritma Naive Bayes dan Support Vector Machine (SVM). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, *4*(2), 375–384. https://doi.org/10.57152/malcom.v4i2.1206

Nur Afrilia, M., Rahaningsih, N., Danar Dana, R., & Dienwati Nuris, N. (2024). Optimasi Analisis Clustering Untuk Aktivitas Dan Respon Pengguna Media Sosial Dengan K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(1), 148–155. https://doi.org/10.36040/jati.v8i1.8334

NURHASANAH, Y. I., & NAUFAL, M. T. (2024). Klasifikasi Sentimen Pada Buzzer Politik Jelang Pemilu 2024 menggunakan Metode Lexicon-based. *MIND Journal*, *9*(2), 166–178. https://doi.org/10.26760/mindjournal.v9i2.166-178

Rahayu, A. S., Fauzi, A., & Rahmat, R. (2022). Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify. *Jurnal Sistem Komputer Dan Informatika (JSON)*, *4*(2), 349. https://doi.org/10.30865/json.v4i2.5398

Rahman, A., & Bustamam, A. (2022). Deep learning with concatenate model to detect COVID-19

lung disease with CT scan images. *AIP Conference Proceedings*, *2391*(1). https://doi.org/10.1063/5.0072411

Rahman, M. S., Ghosh, T., Aurna, N. F., Kaiser, M. S., Anannya, M., & Hosen, A. S. M. S. (2023). Machine learning and internet of things in industry 4.0: A review. *Measurement: Sensors*, *28*(January). https://doi.org/10.1016/j.measen.2023.100822

Rina Noviana, & Isram Rasal. (2023). Penerapan Algoritma Naive Bayes Dan Svm Untuk Analisis Sentimen Boy Band Bts Pada Media Sosial Twitter. *Jurnal Teknik Dan Science*, *2*(2), 51–60. https://doi.org/10.56127/jts.v2i2.791

Siringoringo, R. (2018). KLASIFIKASI DATA TIDAK SEIMBANG MENGGUNAKAN ALGORITMA SMOTE DAN k-NEAREST NEIGHBOR. *Jurnal ISD*, *3*(1), 44–49.

Siti Aisah, I., Irawan, B., & Suprapti, T. (2024). Algoritma Support Vector Machine (Svm) Untuk Analisis Sentimen Ulasan Aplikasi Al Qur'an Digital. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(6), 3759–3765. https://doi.org/10.36040/jati.v7i6.8263

Soraya, F. A., & Indriyanti, A. D. (2024). *Perbandingan Algoritma Naïve Bayes dan Support Vector Machine dalam Analisis Sentimen Aplikasi Teman Bus*. *05*(02), 118–125.

Sulistiyowati, N., & Jajuli, M. (2020). Integrasi Naive Bayes Dengan Teknik Sampling Smote Untuk Menangani Data Tidak Seimbang. *Nuansa Informatika*, *14*(1), 34. https://doi.org/10.25134/nuansa.v14i1.2411

Tarwoto, Nugroho, R., Azka, N., & Graha, W. S. R. (2025). Analisis Sentimen Ulasan Aplikasi Mobile JKN di Google PlayStore Menggunakan IndoBERT. *Jurnal JTIK (Jurnal Teknologi Informasi Dan Komunikasi)*, *9*(2), 495–505. https://doi.org/10.35870/jtik.v9i2.3340

Vincent, R., Maulana, I., & Komarudin, O. (2024). Perbandingan Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Analisis Sentimen Dengan Multiclass Di Twitter. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(4), 2496–2505. https://doi.org/10.36040/jati.v7i4.7152

Widodo, J., & Kusnan, K. (2023). Mapping and Analyzing E-Government Sentiments in Local Governments in Indonesia. Society, 11(2), 434–457. https://doi.org/10.33019/society.v11i2.558

Yulian, T., & Susanto, E. R. (2025). *Sistemasi: Jurnal Sistem Informasi Analisis Perbandingan Teknik Oversampling dan SMOTEENN pada Algoritma Machine Learning untuk Prediksi Penyakit Kanker Payudara Comparative Analysis of Oversampling and SMOTEENN Techniques in Machine Learning Algorithms f*. *14*, 1318–1331. http://sistemasi.ftik.unisi.ac.id