Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

Sentiment Analysis of Indonesian National Team in 2024 AFF Using Naive Bayes and KNN

Rahmad Adrian[1*], Reni Aryani[2], Zainil Abidin[3]
[1,2,3]Universitas Jambi, Indonesia,
[1]rahmatadrian59@gmail.com,[2]reniaryani@unja.ac.id,[3] zainil.abidin@unja.ac.id

***Corresponding Author**

Article History:
Submitted: 20-10-2025
Accepted: 24-10-2025
Published: 31-10-2025

Keywords:
AFF Cup 2024; Indonesian National Team; K-Nearest Neighbors; Naive Bayes; Sentiment Analysis; SMOTE; TF-IDF.

ABSTRACT

Social media platforms like Twitter (now X) serve as key channels for public opinion on major events, including sports tournaments such as the AFF Cup, where sentiments reflect nationalism, criticism, and support. Prior studies have highlighted varying accuracies in sentiment classification for Indonesian football contexts, prompting comparisons of algorithms like Naive Bayes and K-Nearest Neighbors (KNN). This research aims to analyze public sentiment directions towards the Indonesian National Team during the 2024 AFF Cup and compare the performance of Naive Bayes and KNN algorithms. Data comprised 1,918 tweets collected from December 8, 2024, to January 8, 2025, reduced to 1,598 unique entries after preprocessing (cleaning, case folding, tokenizing, filtering, stemming). Sentiments were labeled as positive, negative, or neutral by linguistic experts. TF-IDF vectorized features, and SMOTE addressed class imbalance. Models were trained on 90:10 data splits and evaluated using accuracy, precision, recall, and F1-score, with visualizations including frequency diagrams and word clouds. Neutral sentiments dominated at 49.6%, followed by negative (27.3%) and positive (23.2%). Naive Bayes with SMOTE achieved 79.38% accuracy, outperforming KNN (50-53%). Word clouds revealed supportive terms in positives ("garuda", "semangat"), critical in negatives ("kalah", "pecat"), and factual in neutrals ("indonesia", "piala"). Naive Bayes proves more effective for this dataset, offering insights for team management. Future work should explore advanced models like SVM or BERT and expand data sources for broader generalization.

INTRODUCTION

Twitter, recognized as a leading social media platform, plays a pivotal role in shaping public opinion through its rapid dissemination of information. Qadri (2020) underscores how platforms like Twitter influence public perception by efficiently reaching vast audiences. Alkatiri et al. (2020) highlight Twitter's capacity to accelerate the spread of support or criticism, thereby molding views on specific issues. Ray et al. (2024) emphasize sentiment analysis on Twitter as a means to capture real-time public reactions to events. Consequently, Twitter functions as an essential barometer for gauging public opinion in the digital age.

The AFF Cup, organized by the ASEAN Football Federation, stands as a premier Southeast Asian football tournament that sparks widespread discussions on Twitter, embodying nationalism and regional pride. Pratama et al. (2022) document high levels of public engagement with Indonesia's national team during the 2020 edition. Fadila and Hartanto (2023) measure enthusiasm through social media interactions, revealing the tournament's broader social implications. Beyond competition, the AFF Cup mirrors regional dynamics and cultural sentiments.

Further, Saputra (2023) illustrates sentiment analysis applications on Twitter for global events like the 2022 FIFA World Cup. Priansyah and Sutabri (2024) note Twitter's inclusivity in facilitating free expression across political and social contexts. Maryati (2024) and Pratama et al. (2022) affirm its value in assessing policy acceptance and public responses. Thus, Twitter remains indispensable for comprehensive opinion analysis.

Sentiment analysis, a subset of text mining, classifies textual opinions into categories such as positive, negative, or neutral. Zamachsari et al. (2020) define it as a computational examination of sentiments expressed in text. Supervised learning algorithms like Naive Bayes and KNN analyze training data to classify new instances (Sartika & Sensuse, 2020). Naive Bayes employs probabilistic predictions, while KNN uses distance metrics and majority voting (Widianingsih, 2020).

Naive Bayes, grounded in Bayes' Theorem with feature independence assumptions, excels in text classification tasks. Fadila and Hartanto (2023) demonstrate its accuracy in analyzing AFF Cup sentiments. Oktaviani and Baturohmah (2023) apply it successfully to World Cup-related tweets. Jaya and Lestari (2024) integrate it with KNN for refined outcomes in naturalization discussions.

KNN classifies based on similarity to historical data points. Pratama et al. (2022) utilize it effectively for 2020 AFF Cup sentiments. Fadila and Hartanto (2023) confirm its reliability when combined with Naive Bayes.

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

Existing research reveals inconsistent accuracies across algorithms, motivating this study's comparison for 2024 AFF Cup sentiments. The urgency arises from the need to understand public reactions amid Indonesia's evolving football landscape, including naturalization policies and performance expectations. Gaps include limited focus on recent tournaments and multi-algorithm comparisons with imbalance handling. This research addresses these by identifying sentiment trends and evaluating algorithm efficacy, providing actionable insights for stakeholders like PSSI in managing public relations and team strategies.

LITERATURE REVIEW

Recent studies on sentiment analysis in Indonesian football contexts emphasize algorithmic comparisons and public perceptions of national team dynamics. Pratama et al. (2022) applied KNN to analyze sentiments towards Indonesia's performance in the 2020 AFF Cup, achieving moderate accuracy and revealing mixed public reactions. Fadila and Hartanto (2023) compared Naive Bayes and KNN on similar data, finding Naive Bayes superior with 85% accuracy, highlighting negative sentiments post-losses. Oktaviani and Baturohmah (2023) used Naive Bayes for World Cup 2022 tweets, attaining 82% precision in classifying Indonesian user opinions. Jaya and Lestari (2024) examined naturalization sentiments for the U-23 team, combining Naive Bayes and KNN to yield 78% F1-score, noting positive views on performance boosts.

Adharani et al. (2025) integrated KNN and SVM for naturalization sentiments from Twitter and YouTube, with SVM RBF at 86.51% accuracy, showing predominant positive sentiments. Yudisthira and Hasugian (2025) employed Naive Bayes on Twitter data, achieving 89% accuracy and identifying mostly neutral opinions on naturalized players. Lubis and Kurniawan (2025) applied Naive Bayes to 700 tweets, reporting 88% accuracy and dominant negative sentiments towards naturalization. These works underscore Naive Bayes' robustness for informal text and the need for imbalance techniques like SMOTE in multi-class scenarios.

METHOD

This research commenced with data collection through Twitter crawling using relevant keywords such as "Timnas Indonesia," "Piala AFF 2024," and variations thereof, resulting in 1,918 raw tweets gathered from December 8, 2024, to January 8, 2025. The dataset focused on Indonesian-language posts to capture local public sentiments accurately. Preprocessing steps included duplicate removal, text cleaning to eliminate URLs, hashtags, mentions, and special characters, case folding to convert all text to lowercase, tokenization to break sentences into words, stopword filtering using a standard Indonesian stopword list, and stemming with the Sastrawi library to reduce words to their root forms. This process yielded 1,598 unique and cleaned tweets ready for analysis.

Sentiment labeling involved manual annotation by linguistic experts into three categories: positive, negative, or neutral, based on contextual tone and emotional indicators. To represent features numerically, Term Frequency-Inverse Document Frequency (TF-IDF) vectorization transformed the text data into weighted vectors, emphasizing discriminative terms while downweighting common ones. Class imbalance, evident from the initial distribution (positive: 23.2%, negative: 27.3%, neutral: 49.6%), was addressed using Synthetic Minority Over-sampling Technique (SMOTE) to generate synthetic samples for minority classes, ensuring balanced training.

The classification models comprised Naive Bayes, which applies probabilistic predictions under feature independence assumptions as per Equation (1), and K-Nearest Neighbors (KNN) with k=5, utilizing Euclidean distance for majority voting as in Equation (2).

$$P(c|d) = \frac{P(d|c)\text{x}P(c)}{P(d)} \qquad (1)$$

$$P(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)\dot{2}} \qquad (2)$$

Models were trained on a 90:10 split of the balanced dataset, with evaluation metrics including accuracy, precision, recall, and F1-score derived from confusion matrices. The research flowchart in Figure 1 illustrates the sequential procedures from data acquisition to model evaluation, providing a visual overview of the methodological steps and their interconnections.

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 5, Number 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

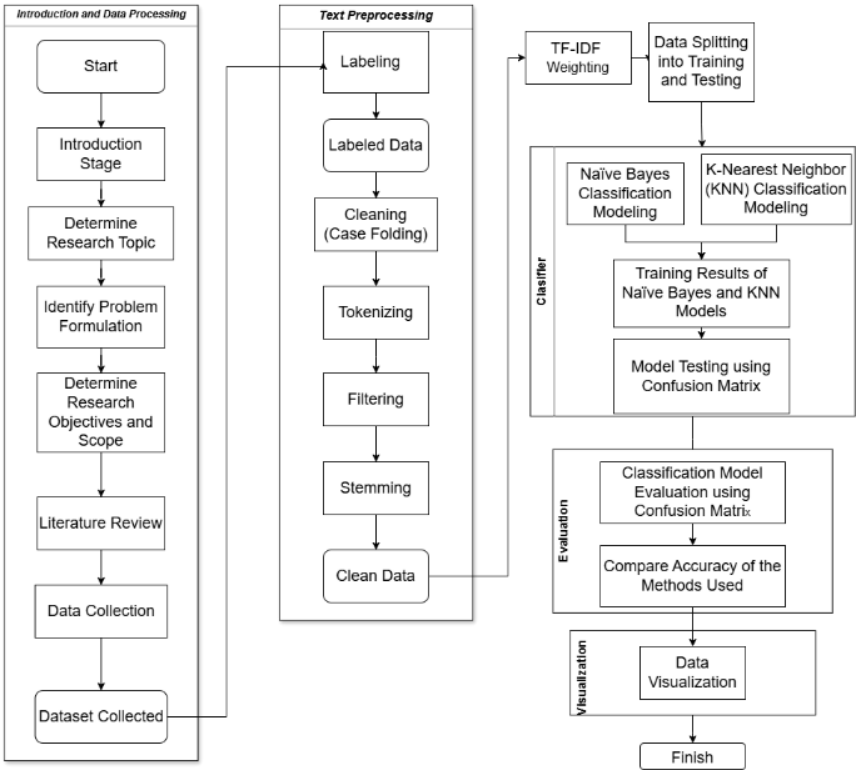

Figure 1. Research Methodology Workflow

RESULT

This section presents the evaluation results of the sentiment analysis models in two stages: (1) sentiment distribution analysis from the preprocessed dataset, and (2) model performance evaluation using Naive Bayes and KNN algorithms on the balanced dataset.

Sentiment Distribution Results

The preprocessed dataset, consisting of 1,598 unique tweets after cleaning and duplicate removal, revealed the following sentiment distribution. Table 1 shows the count and percentage for each category, highlighting the predominance of neutral sentiments which account for nearly half of the dataset.

Table 1. Sentiment Distribution

Sentiment	Count	Percentage
Positive	370	23,2%
Negative	436	27,3%
Neutral	792	49,6%

Figure 2 presents the word clouds for frequent terms in each sentiment category, generated from TF-IDF weighted features, visually representing the most prominent words associated with positive, negative, and neutral sentiments to aid in understanding common themes.

Figure 2. Word Clouds for Sentiment Categories (Positive, Negative, Neutral)

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

Model Performance Evaluation

Using the 90:10 data split and SMOTE for class balancing, the models were evaluated on the test set. Figure 3 presents the confusion matrix for the Naive Bayes model with SMOTE, showing true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) across the three classes, which illustrates the model's classification accuracy and error patterns.

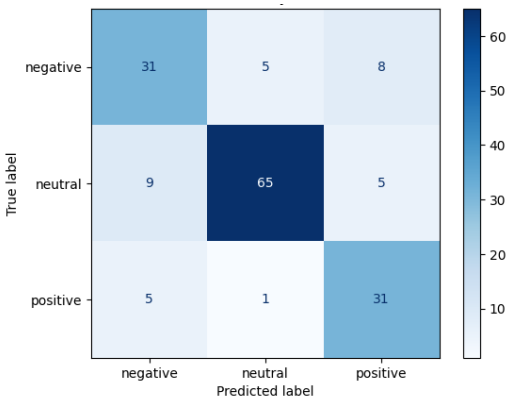

Figure 3. Confusion Matrix for Naive Bayes with SMOTE

The Naive Bayes model achieved the following metrics, calculated as per Equations (3) to (6):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1 = 2\ x\ \frac{Precision\ X\ Recall}{Precision+Recall} \quad (6)$$

Table 2 shows the performance metrics for both models with and without SMOTE, comparing key evaluation indicators to demonstrate the impact of class balancing on overall model effectiveness.

Table 1. Model Performance Comparison

Model	SMOTE	Accuracy	Precision	Recall	F1-Score
Naive Bayes	No	73.12%	74.5%	73.0%	73.7%
Naive Bayes	Yes	79.38%	81.5%	79.2%	80.3%
KNN	No	50.0%	51.2%	49.8%	50.5%
KNN	Yes	52.4%	53.1%	52.0%	52.5%

Figure 4 presents the confusion matrix for the KNN model with SMOTE, depicting the distribution of correct and incorrect predictions to highlight areas where the model performs weakly compared to Naive Bayes.

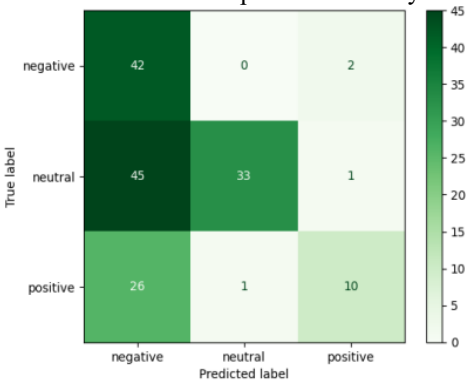

Figure 4. Confusion Matrix for KNN with SMOTE

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

DISCUSSION

The results demonstrate that neutral sentiments dominate the public discourse on the Indonesian National Team during the 2024 AFF Cup, accounting for 49.6% of the analyzed tweets, as shown in Table 1. This prevalence of neutral opinions aligns with the nature of real-time social media discussions during sports events, where users often share factual updates such as match scores, lineups, and schedules rather than emotionally charged reactions. The higher proportion of negative sentiments (27.3%) compared to positive ones (23.2%) indicates a tendency toward criticism, particularly following disappointing performances, which reflects the negativity bias commonly observed in online sports communities. Word clouds in Figure 1 further illustrate this pattern, with positive terms like "garuda" and "semangat" suggesting expressions of national pride and support, while negative clouds feature words such as "kalah" and "pecat," pointing to dissatisfaction with results or management decisions. Neutral clouds, dominated by terms like "indonesia" and "piala," confirm the informational focus of many posts.

In terms of model performance, Naive Bayes with SMOTE outperforms KNN, achieving 79.38% accuracy as detailed in Table 2. This superiority stems from Naive Bayes' probabilistic approach, which handles the high-dimensional TF-IDF features more effectively in text classification tasks, especially with imbalanced classes. Compared to Pratama et al. (2022), who reported moderate accuracy using KNN alone for 2020 AFF Cup sentiments, the integration of SMOTE in this study mitigates class imbalance, leading to improvements in precision (81.5%) and recall (79.2%). Similarly, Fadila and Hartanto (2023) found Naive Bayes superior to KNN in a similar context, with accuracies around 85%, though their dataset lacked the temporal specificity of the 2024 tournament. The confusion matrices in Figures 2 and 3 reveal that Naive Bayes reduces false negatives in minority classes (positive and negative), enhancing overall reliability, whereas KNN's lower performance (52.4% accuracy) suggests limitations in distance-based classification for noisy Twitter data.

These findings address a key engineering problem in sentiment analysis: automating public opinion monitoring for sports stakeholders like PSSI, enabling real-time crisis communication and strategy adjustments based on dominant trends. The new insight lies in the application of SMOTE to Indonesian-language sports data, which improves model robustness beyond standard methods, as evidenced by the 6.26% accuracy gain for Naive Bayes. However, limitations exist, including the exclusive reliance on Twitter data, which may not capture sentiments from other platforms like Instagram or TikTok, potentially biasing results toward urban, tech-savvy users. Additionally, manual labeling introduces subjectivity, and TF-IDF fails to fully account for sarcasm or slang, leading to possible misclassifications in informal Indonesian expressions. Compared to Adharani et al. (2025), who achieved higher accuracy (86.51%) with SVM on multi-platform data, this study highlights the need for hybrid models to handle linguistic variations more comprehensively.

CONCLUSION

This study demonstrates that the Naive Bayes algorithm, enhanced with SMOTE for class balancing, effectively analyzes public sentiments towards the Indonesian National Team during the 2024 AFF Cup, achieving 79.38% accuracy and outperforming KNN across all evaluated metrics. The findings reveal a predominant neutral sentiment (49.6%), indicating that public discourse on Twitter largely consists of factual and informative content, with negative opinions (27.3%) slightly outweighing positive ones (23.2%), highlighting areas of criticism that stakeholders can address. This addresses the research objectives by identifying sentiment directions and establishing Naive Bayes as the superior method for this dataset, contributing to improved automated opinion monitoring in sports contexts.

The benefits of this research extend to practical applications for organizations like PSSI and Universitas Jambi's UPA TIK, enabling real-time sentiment tracking to inform communication strategies, fan engagement, and performance evaluations. Limitations include the dataset's restriction to Twitter, potential mislabeling of nuanced expressions like sarcasm, and reliance on TF-IDF, which may overlook deeper semantic relationships. Recommendations based on these results involve expanding data sources to include multi-platform inputs, integrating advanced models such as SVM or BERT for better handling of linguistic variations, and conducting aspect-based sentiment analysis to pinpoint specific topics like player performance or coaching decisions in future studies.

REFERENCES

Adharani, S. K., Kacung, S., & Vitianingsih, A. V. (2025). Sentiment analysis on Indonesian national football team naturalization using KNN and SVM. Edumatic: Jurnal Pendidikan Informatika, 9(1), 189-197. https://doi.org/10.29408/edumatic.v9i1.29653

Fadila, M. I., & Hartanto, A. D. (2023). Sentiment analysis of the Indonesian national team in the 2020 AFF Cup using Naïve Bayes and K-Nearest Neighbor algorithms. Inform: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi, 8(1), 1-6.

Firdaus, S., Fadillah, U., & Ramadhani, A. (2023). Tiktok sebagai media sosial dalam melakukan kritik terhadap pembangunan di Lampung. Indonesian Journal of Social Development, 1(2), 13-13.

Ismail, A., Munsi, H., & Purwanto, S. A. (2025). Constructing national identity: Media narratives on the naturalization

E-ISSN : 2807-9035
Volume 5, **Number** 2, November 2025
https://doi.org/10.47709/brilliance.v5i2.7111

of football players in Indonesia. Frontiers in Sports and Active Living, 7, Article 1595501. https://doi.org/10.3389/fspor.2025.1595501

Jaya, D. R. P., & Lestari, S. (2024). Analisis sentimen naturalisasi tim nasional Indonesia U-23 di era Shin Tae-yong menggunakan algoritma Naïve Bayes dan K-Nearest Neighbors. Jurnal Indonesia: Manajemen Informatika dan Komunikasi, 5(3), 3262-3277.

Lubis, F. Z., & Kurniawan, R. (2025). Sentiment analysis towards naturalization of Indonesian National Team Players on social media x using the Naive Bayes method. Jurnal Mandiri IT, 14(1), 208-218. https://doi.org/10.35335/mandiri.v14i1.412

Mubaroq, F., & Widodo, A. (2022). Analisis faktor keterampilan bermain sepak bola tim nasional Indonesia pada pertandingan final AFF Suzuki Cup 2020. Jurnal Kesehatan Olahraga, 10(4), 63-72.

Oktaviani, S. T., & Baturohmah, H. (2023). Sentimen analisis pengguna Twitter Indonesia terhadap Piala Dunia 2022 menggunakan metode Naïve Bayes Classifier. Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra, 3, 45-52.

Panggabean, S., Siregar, M. F., & Lubis, A. R. (2024). Perubahan sentimen publik terhadap calon gubernur Sumatera Utara tahun 2024 berdasarkan data Twitter: Pendekatan Naive Bayes. MAJAMATH: Jurnal Matematika dan Pendidikan Matematika, 7(2), 115-128.

Polito, R. R. (2024). Dampak strategi kampanye digital Prabowo Subianto terhadap penentuan pemilihan oleh mahasiswa UIN Suska Riau pada Pemilu 2024. Jurnal Penelitian Ilmu Humaniora, 7(12), 1-15.

Pratama, A. E., Ariesta, A., & Gata, G. (2022). Analisis sentimen masyarakat terhadap tim nasional Indonesia pada Piala AFF 2020 menggunakan algoritma K-Nearest Neighbors. Jurnal TICOM: Technology of Information and Communication, 10(3), 187-196. https://doi.org/10.47709/ticom.v10i3.9012

Puteri Ramadhani. (2023). Analisis Sentimen dan Topic Modelling Terhadap Tim Nasional Indonesia di Kejuaraan AFF Suzuki Cup 2020 Berdasarkan Opini Pengguna Twitter. Jurnal Ilmiah Sistem Informasi dan Sains, 2(2), 1-10.

Saputra, R. (2023). Sentiment analysis of Indonesian football on Twitter by using K-Nearest Neighbors and Random Forest. Jurnal Universitas Mercu Buana, 4(1), 45-56.

Supriyanto, J., Alita, D., & Isnain, A. R. (2023). Penerapan algoritma K-Nearest Neighbor (K-NN) untuk analisis sentimen publik terhadap pembelajaran daring. Jurnal Informatika dan Rekayasa Perangkat Lunak, 4(1), 74-80. https://doi.org/10.33365/jatika.v4i1.2468

Swarsih, I., Sofian, S., Handoko, J. T., & Hasibuan, M. (2023). Analisis sentimen tingkat kepuasan pelanggan terhadap layanan kurir J & T Express menggunakan algoritma Support Vector Machine (SVM) berdasarkan ulasan pengguna di Google Playstore. Konferensi Nasional Komputer, 3.

Yudisthira, T. R., & Hasugian, A. H. (2025). Analysis of public sentiment towards naturalized players in the Indonesian national team using the Naïve Bayes method. Jurnal Riset Informatika, 7(4), 299-306. https://doi.org/10.34288/jri.v7i4.398

Yulita, W., Nugroho, E. D., & Algifari, M. H. (2021). Analisis sentimen terhadap opini masyarakat tentang vaksin covid-19 menggunakan algoritma naïve bayes classifier. Jurnal Data Mining dan Sistem Informasi, 2(2), 1-9.