Jurnal Mandiri IT
Vol. 15 No. 1, July (2026), pp. 19-26

ISSN 2301-8984 (Print), 2809-1884 (Online)
Published by Institute of Computer Science

Sentiment analysis of public opinion on the non-cash food assistance (BPNT) program on platform x using naive bayes

Nurul Asma As-shidiq[1], Donny Maulana[2], M. Zubair Abdurrohman[3]
[1,2,3]Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Bekasi, Indonesia

ARTICLE INFO

Article history:

Received Jun 20, 2026
Revised Jun 27, 2026
Accepted Jul 5, 2026

Keywords:

BPNT;
Naive Bayes;
Platform X;
Sentiment Analysis;
TF-IDF.

ABSTRACT

The Non-Cash Food Assistance (BPNT) program is one of the Indonesian government's social assistance policies that frequently receives public attention and discussion on social media. This study aims to analyze public sentiment toward the implementation of the BPNT program based on comments posted on Platform X. The dataset consisted of 1,758 Indonesian-language comments collected through web crawling between October 2025 and January 2026. The collected data were processed through several preprocessing stages, including case folding, cleaning, tokenization, normalization, stopword removal, and stemming. Furthermore, TF-IDF was applied to transform textual data into numerical features, and sentiment classification was performed using the Multinomial Naive Bayes algorithm. The dataset was divided into training and testing data using an 80:20 ratio. The results showed that neutral sentiment dominated public discussions with 51.48%, followed by positive sentiment with 33.90% and negative sentiment with 14.62%. Performance evaluation using a Confusion Matrix achieved an accuracy of 79.545%. These findings indicate that the Naive Bayes approach can effectively classify public sentiment regarding the BPNT program and provide useful insights for evaluating social assistance policies.

Corresponding Author:

Nurul Asma As-shidiq,
Program Studi Teknik Informatika,
Universitas Pelita Bangsa,
Jl. Inspeksi Kalimalang No.9, Cibatu, Cikarang Sel., Kabupaten Bekasi, Jawa Barat, 17530, Indonesia
Email: nurul13@mhs.pelitabangsa.ac.id

1. INTRODUCTION

The rapid development of information technology has brought fundamental changes in how society interacts and expresses opinions in the public sphere. Social media functions not only as a personal communication tool but also as a digital public sphere that reflects public dynamics towards various social, economic, and government policies. Platform X (formerly Twitter) is one of the microblogging platforms actively used to voice public opinion due to its open, fast, and short-text nature, making it a potential data source for understanding public perceptions in real-time (Damanik & Setyohadi, 2021). In modern governance, public opinion has a strategic role; hence, evaluations should not only rely on conventional administrative reports or surveys but also consider digital public aspirations to support data-driven decision-making (Damanik & Setyohadi, 2021; Febrilliani & Wibowo, 2025).

One social assistance policy receiving significant public attention in Indonesia is the Non-Cash Food Assistance (Bantuan Pangan Non Tunai / BPNT) program. Designed to replace conventional physical food aid with an electronic non-cash mechanism, the BPNT program aims to increase transparency, accountability, and targeting accuracy. However, its implementation in the field still faces various structural challenges. Issues such as inaccurate recipient data, sub-optimal food quality, and distribution technicalities often spark public complaints (Salam dkk., 2023). These problems trigger various public responses on Platform X, ranging from criticisms to support.

Consequently, understanding these responses becomes essential for evaluating the effectiveness and transparency of the BPNT program. Sentiment analysis contributes to improving transparency by providing an independent, data-driven evaluation mechanism that captures direct public feedback, enabling policymakers to identify distribution issues and improve accountability in social assistance programs.

Evaluating this massive volume of unstructured text data manually is highly inefficient and prone to subjectivity. Therefore, a computational text-mining approach is required to systematically process large-scale text data (Salam dkk., 2023). Sentiment analysis, a subfield of Natural Language Processing (NLP), can classify these text opinions into specific polarities positive, negative, and neutral based on word patterns and linguistic context (Damanik & Setyohadi, 2021; Naraswati dkk., 2021). To achieve a high performance on Indonesian-language text, this computational process requires effective algorithms. The Multinomial Naive Bayes algorithm is widely selected for text classification because of its simple structure, computational efficiency, and high competitiveness on short social media texts (Setyawan dkk., 2023; Zulfikar dkk., 2023). Furthermore, optimizing a lightweight computational architecture is highly crucial for maintaining system responsiveness when processing dynamic digital data volumes (Maulana & Ramasamy, 2026).

Several recent studies have leveraged Naive Bayes to capture public sentiment regarding government programs, demonstrating stable and consistent classification performance (Naraswati dkk., 2021; Pristiyono dkk., 2021; Zulfikar dkk., 2023). Other studies have also implemented feature selection techniques to eliminate irrelevant features, significantly boosting the accuracy of the model (Rachmad dkk., 2022), or used sentiment distributions to measure public trust in government agencies (Negara dkk., 2024; Zakaria dkk., 2023). While previous research heavily gravitated towards public health emergencies (such as COVID-19) or energy sectors (fuel price hikes) (Damanik & Setyohadi, 2021; Pristiyono dkk., 2021; Zakaria dkk., 2023), this study specifically isolates and addresses a critical social safety net domain: the food security policy through the BPNT program.

To provide a well-focused analysis, this study sets specific constraints. The data acquisition is restricted to Indonesian public comments on Platform X containing keywords such as “BPNT”, “Bantuan Pangan”, “Saldo BPNT”, and “Kartu Sembako” collected during a crucial initial distribution window from October 1, 2025, to January 25, 2026. The text data is transformed into numerical vectors using Term Frequency Inverse Document Frequency (TF-IDF) weighting and classified using the Multinomial Naive Bayes classifier via Python implementation (specifically Scikit-learn and the Sastrawi library for Indonesian stemming). After automated cleaning and rigorous deduplication processes, a total of 1,758 unique clean data rows are structured for the classification pipeline.

Ultimately, this research aims to implement a systematic NLP preprocessing pipeline for Indonesian text, evaluate the classification performance of Naive Bayes through a Confusion Matrix (accuracy, precision, recall, and F1-score), map the dominant public sentiment distribution regarding the BPNT program, and discuss its practical implications for enhancing the future efficacy of social assistance policies. However, evaluating the BPNT program through a digital data-driven approach possesses inherent limitations, particularly demographic bias, as the dataset predominantly reflects opinions from tech-savvy and urban internet users, potentially underrepresenting rural beneficiaries who lack digital access, thereby limiting the generalizability of the findings.

2. RESEARCH METHOD

The methodological framework applied in this study is structured into systematic stages to process, classify, and evaluate social media public sentiment toward the Non-Cash Food Assistance (BPNT) program. The operational architecture encompasses data acquisition, textual preprocessing, manual annotation, feature extraction, probabilistic classification, and model evaluation.

Data Acquisition and Parameters

Textual data was acquired through automated data crawling on Platform X using Python, Google Colab, and the Pandas library. The data harvesting process targeted public comments written in the Indonesian language over a continuous four-month period spanning from October 1, 2025, to January 25, 2026. Queries were constrained utilizing explicit program-related keywords:

"BPNT", "Bantuan Pangan", and "Saldo BPNT". Following rigorous screening, manual cleaning, and deduplication processes to remove redundant entries, a finalized corpus consisting of 1,758 unique clean rows of public comments was structured for the classification pipeline.

Text Preprocessing Pipeline

Since the raw comments harvested from social media exhibit high text noise, grammatical irregularities, and colloquialisms, a sequence of text-mining preprocessing operations was enforced to normalize the dataset, beginning with case folding to convert all alphabetical characters within the text into a uniform lowercase format, thereby eliminating token variation induced by arbitrary capitalization. This was followed by a cleaning process to truncate metadata noise and irrelevant components including URLs, platform usernames, hashtags, numerical values, punctuation marks, and redundant white spaces. Next, tokenizing was executed to split continuous text strings into individual tokens or standalone word units based on blank space delimiters for isolated analysis (Bird dkk., 2009).

The tokens then underwent a normalization phase to transform slang, typos, non-standard terms, and abbreviations into their standardized Indonesian forms. This normalization process employed a manually curated Indonesian slang and abbreviation dictionary to convert informal expressions, abbreviations, and common spelling variations into their formal Indonesian equivalents before stemming. Examples include converting colloquial forms such as "gk" into "tidak", "yg" into "yang", "dr" into "dari", and "bgt" into "banget".

Afterward, stopword removal was applied to filter out high-frequency words that lack significant semantic sentiment weight, such as conjunctions and prepositions (Bird dkk., 2009). Finally, stemming was performed to strip derivational affixes from words to retrieve their core root forms using the specialized Sastrawi library for Indonesian language processing (Prasetyo & Hermawan, 2023).

Sentiment Labeling

The sentiment labels were assigned manually based on the semantic polarity of each comment by considering its emotional tone, contextual meaning, and expressed opinion toward the implementation of the BPNT program. Positive sentiment explicitly expresses public support, satisfaction, or appreciation regarding the implementation of the BPNT program. Conversely, negative sentiment encompasses public complaints, criticisms, technical issues, or grievances regarding aid distribution. Lastly, neutral sentiment represents purely objective information, factual statements, or queries that do not convey a distinct emotional stance toward the policy.

Feature Extraction via TF-IDF

To translate qualitative textual content into an optimal numerical format for the machine learning algorithm, the Term Frequency-Inverse Document Frequency (TF-IDF) vectorization technique was applied (Manning dkk., 2009; Salton & Buckley, 1988). This method calculates a statistical weight reflection based on term localized frequency (TF) balanced against its dispersion across the entire 1,758 document corpus (IDF). The extraction parameters were constrained to an unigram configuration (ngram_range=(1,1)) with a maximum vocabulary ceiling set at 500 features (max_features=500).

Classification and Model Validation

The probabilistic text classification was conducted using the Multinomial Naive Bayes classifier, which is chosen for its light computational architecture and high efficiency in processing high-dimensional text vectors from social media (Suryani dkk., 2023). Model validation was managed via the hold-out evaluation method, utilizing a fixed 80:20 distribution split. This ratio was selected because it is widely adopted in supervised machine learning studies, as it provides sufficient training data for model learning while preserving an independent testing set for reliable performance evaluation and generalization.

Out of the 1,758 entries, 1,406 documents were allocated as the training dataset to calculate prior and likelihood probabilities. The remaining 352 documents were retained as an independent testing dataset to evaluate the predictive model. For smoothing hidden zero-probabilities of unobserved features during testing, Laplace smoothing was parameterized at $\alpha = 0{,}8$ (Suryani dkk., 2023).

Experimental Setup and Evaluation Metrics

The computational environment was powered by an Intel Core i7 processor laptop alongside cloud-based infrastructure in Google Colab utilizing Python 3.12. Core data operations and NLP steps relied on specialized libraries including Scikit-learn, Pandas, NLTK, Sastrawi, and Tweet Harvest (Pedregosa dkk., 2011).

The classification performance was comprehensively quantified through a 3×3 Confusion Matrix to track classification match rates across the positive, negative, and neutral classes (Normawati & Prayogi, 2021). The model's reliability was mathematically measured using standard metrics (Fawcett, 2006; Powers, 2020):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

To finalize the pipeline, the outputs are visually mapped at 300 DPI high-resolution using Matplotlib and Seaborn to produce Wordclouds for each sentiment class, a Confusion Matrix heatmap, distribution pie charts, and precision-recall evaluation graphs.

3. RESULTS AND DISCUSSIONS

Sentiment Distribution Analysis

The dataset containing 1,758 unique public comments regarding the Non-Cash Food Assistance (BPNT) program was categorized into three distinct sentiment classes using an automated lexicon-based approach tailored to social media syntax (Naraswati dkk., 2021). The structured distribution of public sentiment polarities is presented in table 1.

Table 1. Sentiment distribution of the BPNT dataset

Sentiment Class	Total Document	Percentage (%)
Positive Sentiment	596	33.90%
Negative Sentiment	257	14.62%
Neutral Sentiment	905	51.48%
Dataset Corpus	1,758	100.00%

As outlined in table 1 neutral sentiment constitutes the largest portion of the corpus, accounting for 51.48% of the total dataset (Naraswati dkk., 2021). This clear dominance demonstrates that the majority of public interactions on Platform X are characterized by objective inquiries, factual discussions regarding distribution timelines, official ministry announcements, and administrative procedures regarding card balance checks (Naraswati dkk., 2021). Positive sentiment forms the second-largest share at 33.90%, driven primarily by expressions of public appreciation, gratitude ("alhamdulillah"), and statements highlighting the streamlined distribution of funds that alleviate daily household financial burdens (Naraswati dkk., 2021). Conversely, negative sentiment constitutes the smallest share at 14.62% (Naraswati dkk., 2021). These negative expressions stem from citizen grievances, administrative anomalies, occurrences of unauthorized deductions ("pungli"), unmapped recipients, or technical difficulties such as empty card balances during the distribution window (Naraswati dkk., 2021).

The findings also demonstrate the practical value of real-time sentiment analysis for government agencies. By continuously monitoring public responses on Platform X, policymakers can identify emerging complaints, detect distribution anomalies more rapidly, and respond to operational issues before they escalate into broader administrative problems. Compared with conventional survey-based evaluations, this approach provides faster and more timely evidence to support policy monitoring and decision-making.

Furthermore, the proposed framework has the potential to be integrated into a real-time sentiment monitoring dashboard for the BPNT program. Such a system could continuously collect and analyze public feedback from social media, enabling the Ministry of Social Affairs to monitor public sentiment trends and evaluate policy implementation more efficiently.

Lexical Feature Mapping

Textual word frequencies were visually mapped to extract the most dominant terms driving each sentiment class (Naraswati dkk., 2021). The positive corpus is heavily populated by optimistic root verbs and expressions of relief, including 'cair' (disbursed), 'bantu' (assist), 'manfaat' (benefit), 'alhamdulillah' (praise be to God), 'lancar' (smooth), and 'aman' (secure) (Naraswati dkk., 2021). These high-frequency key tokens underscore that a substantial portion of public feedback reflects operational success and appreciation of the program's real-world utility (Naraswati dkk., 2021).

The negative corpus, conversely, showcases pronounced public frustration driven by underlying systemic or localized issues. The prominent lexical tokens appearing in this segment include 'miskin' (poor), 'telat' (delayed), 'lapor' (report), 'gagal' (failed), 'pungli' (illegal levies), 'potong' (deductions), and 'korupsi' (corruption). This highlights that field bottlenecks and administrative deviations are the key drivers behind public dissatisfaction.

These negative sentiment patterns provide valuable evidence for policymakers to improve the aid distribution system. Recurring complaints regarding delayed disbursement, unauthorized deductions, inaccurate recipient data, and empty account balances indicate operational weaknesses that should be prioritized in future program improvements and monitoring activities.

Conversely, the dominance of positive expressions associated with smooth fund disbursement indicates that timely and transparent aid distribution contributes to strengthening public trust in government social assistance programs. As distribution reliability improves, public confidence and acceptance of the BPNT program are also expected to increase.

Lastly, the neutral corpus is strictly characterized by formal, operational, and non-emotional keywords such as 'pangan' (food), 'masyarakat' (society), 'perintah' (government), 'kartu' (card), 'sembako' (groceries), 'kemensos' (Ministry of Social Affairs), 'cek' (check), and 'saldo' (balance).

Empirical Classification and Model Performance

Model validation was strictly scrutinized by mapping the testing predictions against the empirical ground truth annotations via a 3×3 Confusion Matrix (Hermawan dkk., 2024; Naraswati dkk., 2021). The independent testing partition comprised exactly 352 unseen documents (derived from the 80:20 split of the 1,758 total records) (Hermawan dkk., 2024; Naraswati dkk., 2021). The real-world classification match distribution is summarized in table 2.

Table 2. Empirical matrix cross tabulation

Actual \ Predicted	Negative	Neutral	Positive	Total Actual
Negative	29	14	5	48
Neutral	5	152	19	176
Positive	0	29	99	128
Total Predicted	34	195	123	352

Based on the raw data from the Confusion Matrix, the global classification accuracy achieved by the Multinomial Naive Bayes model is mathematically resolved as follows:

$$Accuracy = \frac{\sum True\ Positive}{Total\ Testing\ Documents} = \frac{29 + 152 + 99}{352} = \frac{280}{352} = 79.545\%$$

The empirical data demonstrates that the model successfully classifies 280 out of 352 test items correctly, establishing a solid baseline accuracy of 79.545%. To provide a more granular evaluation of how the classifier behaves within each unique emotional dimension, individual class metrics are detailed in table 3.

Table 3 Class level performance breakdown

Sentiment Class	Precision	Recall	F1-Score	Support (Test Data)
Positive	0.80	0.77	0.79	128 Data
Negative	0.85	0.60	0.71	48 Data
Neutral	0.78	0.86	0.82	176 Data
Global Accuracy		79.545%		352 Data

Discussion and Error Identification

The empirical results prove that the integrated NLP preprocessing pipeline (consisting of case folding, cleaning, tokenizing, normalization, stopword removal, and Sastrawi-driven stemming) serves as an effective mechanism for streamlining unstructured social media texts (Hermawan dkk., 2024; Naraswati dkk., 2021). By successfully stripping noisy components, eliminating redundant symbols, and standardizing irregular colloquialisms into proper base lemmas, the feature space is optimized, allowing the Multinomial Naive Bayes model to detect underlying sentiment patterns with a high degree of confidence (Hermawan dkk., 2024; Naraswati dkk., 2021).

Nevertheless, a focused error analysis reveals specific vulnerabilities in the classification architecture, primarily localized within the negative and neutral spectrums (Naraswati dkk., 2021). As shown in table 2, 14 actual negative comments were misclassified as neutral, and 19 actual neutral statements were incorrectly tagged as positive. This performance friction is caused by three key factors: a) Semantic Overlap and Shared Vocabularies: Many tweets discuss identical operational parameters (such as 'cair', 'bansos', and 'saldo'), but differ entirely in emotional intent. For instance, a neutral query regarding a balance update can easily confuse the classifier due to lexical similarities with negative complaints about empty balances; b) Colloquial Slang and Unstandardized Abbreviations: Despite running a dedicated normalization step, highly dynamic social media shortcuts, non-standard slang variations, and mixed structures occasionally escape the cleaning filters, resulting in out-of-vocabulary anomalies; c) Sarcasm and Contextual Nuance: The Naive Bayes classifier operates on a conditional independence assumption, meaning it analyzes words as isolated entities rather than tracking contextual dependencies. Consequently, when a user employs sarcastic phrasing (e.g., using positive-weighted words to mock a delay in distribution), the model fails to capture the underlying negative tone and misclassifies the comment.

Overall, despite these localized linguistic challenges, the achieved accuracy of 79.55% confirms that the proposed Multinomial Naive Bayes framework remains a highly viable, computationally lightweight, and practical tool for real-time policy evaluation and public feedback tracking (Hermawan dkk., 2024; Naraswati dkk., 2021).

4. CONCLUSION

This study successfully achieved its research objectives by applying TF-IDF and Multinomial Naive Bayes to classify public sentiment toward the BPNT program. As anticipated, implementing a dedicated Natural Language Processing (NLP) pipeline comprising case folding, text cleaning, tokenization, normalization, stopword filtering, and Sastrawi-based stemming effectively streamlined irregular Indonesian public text on Platform X and prepared it for structural mathematical modeling. The vectorization through Term Frequency-Inverse Document Frequency (TF-IDF) successfully mapped qualitative lexical tokens into optimized numerical feature weights, which enabled the Multinomial Naive Bayes algorithm to establish a clear baseline global classification accuracy of 79.545%. The resulting sentiment distribution revealing 51.48% neutral, 33.90% positive, and 14.62% negative classes directly addresses the primary research objective by providing an objective, data-driven map of public reaction toward the Non-Cash Food Assistance (BPNT) program.

The sentiment classification results imply that improving the quality of social assistance services requires a more responsive and data-driven evaluation mechanism. By systematically identifying public complaints, technical distribution issues, and beneficiary feedback, government agencies can prioritize corrective actions, improve service delivery, and strengthen accountability in future BPNT implementation.

The proposed framework can be further developed using more advanced machine learning methods and applied as a real-time public sentiment monitoring system. In terms of technical development, the framework can be scaled by integrating advanced contextual architectures or ensemble machine learning models to effectively decode semantic exceptions, internet slang, and structural irony that typically hinder conditional independence models. In terms of practical application, this lightweight and computationally efficient framework can be adapted into an automated, real-time public dashboard for government monitoring bodies, such as the Ministry of Social Affairs. By shifting policy evaluation from traditional periodic surveys to continuous social media sentiment analysis, relevant stakeholders can immediately detect localized distribution anomalies, empty balance reports, or regional illegal levies as they surface online (Naraswati et al., 2021). This shift ultimately transforms digital public opinion into a practical asset for adaptive, data-driven policy adjustment and improved social safety net transparency.

This research contributes to the development of data-driven public policy by demonstrating that Natural Language Processing and Multinomial Naive Bayes can effectively transform large-scale social media data into actionable evidence for monitoring, evaluating, and continuously improving social assistance programs.

ACKNOWLEDGEMENTS

The author wishes to express sincere gratitude and high appreciation to all parties who provided invaluable academic guidance, technical insights, and continuous support throughout the completion of this research project. Special thanks are extended to the thesis advisors for their patient mentorship, constructive critiques, and dedication to refining the methodological rigor of this study. The author also acknowledges the open-source community for providing the computational tools, libraries, and resources that made the data preprocessing and classification pipelines possible. Finally, profound gratitude goes to family and peers for their endless encouragement and support during every stage of this academic journey.

REFERENCES

Alfawas, T. I., & Rahim, A. (t.t.). Penerapan Fitur Ekstraksi TF-IDF untuk Analisis Sentimen Ulasan Game Bus Simulator Indonesia dengan Algoritma Naive Bayes.

Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. O'Reilly Media, Inc.

Damanik, F. J., & Setyohadi, D. B. (2021). Analysis Of Public Sentiment About Covid-19 In Indonesia On Twitter Using Multinomial Naive Bayes And Support Vector Machine. IOP Conference Series: Earth and Environmental Science, 704(1), 012027. https://doi.org/10.1088/1755-1315/704/1/012027

Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, ROC Analysis in Pattern Recognition, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Febrilliani, J. S., & Wibowo, A. (2025). Twitter Public Sentiment towards Vasectomy in Indonesia Using SVM and Naïve Bayes. Research Horizon, 5(4), 1415–1424. https://doi.org/10.54518/rh.5.4.2025.756

Hermawan, D., Aditya, A., & Purwiantono, F. E. (2024). Sentiment Analysis of Vaccine Perceptions in Indonesia: A Comparative Study Using the Naïve Bayes Multinomial Text Algorithm on Twitter Data. 2024 International Seminar on Intelligent Technology and Its Applications (ISITIA), 42–47. https://doi.org/10.1109/ISITIA63062.2024.10667933

Manning, C., Raghavan, P., & Schuetze, H. (2009). Introduction to Information Retrieval.

Maulana, D., & Ramasamy, R. K. (2026). Systematic Review of Quantization-Optimized Lightweight Transformer Architectures for Real-Time Fruit Ripeness Detection on Edge Devices. Computers, 15(1), 69. https://doi.org/10.3390/computers15010069

Muthmainnah, M., Rudiman, R., & Yulianto, F. (2025). Sentiment Analysis of the Public on the Deployment of Smart Robots in Indonesia Using the Naïve Bayes Method. JSE Journal of Science and Engineering, 3(2), 85–89. https://doi.org/10.30650/jse.v3i2.3887

Naraswati, N. P. G., Nooraeni, R., Rosmilda, D. C., Desinta, D., Khairi, F., & Damaiyanti, R. (2021). Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification. Sistemasi: Jurnal Sistem Informasi, 10(1), 222–238. https://doi.org/10.32520/stmsi.v10i1.1179

Negara, E. S., Syaputra, R., Erlansyah, D., Andryani, R., Saksono, P. H., Aditya, F., & Agam, P. M. (2024). SENTIMENT ANALYSIS OF PUBLIC OPINION ON PRESIDENTIAL ADVISORY APPOINTMENTS USING NAIVE BAYES CLASSIFICATION. Jurnal Ilmiah Ilmu Terapan Universitas Jambi, 8(2), 452–466. https://doi.org/10.22437/jiituj.v8i2.35254

Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. 5.

Nurlita, N., Rudiman, R., & Hallim, A. (2025). Sentiment Analysis, Classification of Twitter Tweets of the Constitutional Court Decision on Vice Presidential Candidates Using Naïve Bayes Method. JSE Journal of Science and Engineering, 3(2), 96–100. https://doi.org/10.30650/jse.v3i2.3894

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12(85), 2825–2830.

Powers, D. M. W. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation (arXiv:2010.16061). arXiv. https://doi.org/10.48550/arXiv.2010.16061

Prasetyo, P. A. M., & Hermawan, A. (2023). Analisis sentimen twitter terhadap pemilihan presiden menggunakan algoritma naïve bayes. INFOTECH : Jurnal Informatika & Teknologi, 4(2), 224–233. https://doi.org/10.37373/infotech.v4i2.863

Pristiyono, Ritonga, M., Ihsan, M. A. A., Anjar, A., & Rambe, F. H. (2021). Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm. IOP Conference Series: Materials Science and Engineering, 1088(1), 012045. https://doi.org/10.1088/1757-899X/1088/1/012045

Rachmad, A., Kustiyahningsih, Y., Pratama, R. I., Syakur, M. A., Rochman, E. M. S., & Hapsari, D. (2022). Sentiment Analysis of Government Policy Management on the Handling of Covid-19 Using Naive Bayes with Feature Selection. 2022 IEEE 8th Information Technology International Seminar (ITIS), 156–161. https://doi.org/10.1109/ITIS57155.2022.10010004

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

Salam, R. R., Jamil, M. F., Ibrahim, Y., Rahmaddeni, R., Soni, S., & Herianto, H. (2023). Analisis Sentimen Terhadap Bantuan Langsung Tunai (BLT) Bahan Bakar Minyak (BBM) Menggunakan Support Vector Machine: Sentiment Analysis of Cash Direct Assistance Distribution for Fuel Oil Using Support Vector Machine. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 3(1), 27–35. https://doi.org/10.57152/malcom.v3i1.590

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. Information Processing & Management, 24(5), 513–523. https://doi.org/10.1016/0306-4573(88)90021-0

Setyawan, H., Azizah, L. M., & Pradani, A. Y. (2023). Sentiment Analysis of Public Responses on Indonesia Government Using Naïve Bayes and Support Vector Machine. Emerging Information Science and Technology, 4(1), 1–7. https://doi.org/10.18196/eist.v4i1.18681

Suryani, S., Fayyad, M. F., Savra, D. T., Kurniawan, V., & Estanto, B. H. (2023). Sentiment Analysis of Towards Electric Cars using Naive Bayes Classifier and Support Vector Machine Algorithm. Public Research Journal of Engineering, Data Technology and Computer Science, 1(1), 1–9. https://doi.org/10.57152/predatecs.v1i1.814

Zakaria, Z., Kusrini, K., & Ariatmanto, D. (2023). Sentiment Analysis to Measure Public Trust in the Government Due to the Increase in Fuel Prices Using Naive Bayes and Support Vector Machine. International Journal of Artificial Intelligence & Robotics (IJAIR), 5(2), 54–62. https://doi.org/10.25139/ijair.v5i2.7167

Zulfikar, W. B., Atmadja, A. R., & Pratama, S. F. (2023). Sentiment Analysis on Social Media Against Public Policy Using Multinomial Naive Bayes. Scientific Journal of Informatics, 10(1), 25–34. https://doi.org/10.15294/sji.v10i1.39952