Jurnal Mandiri IT
Vol. 14 No. 2, October (2025), pp. 170-178

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

Sentiment analysis of privacy issues in the digital era using the naïve bayes method

Rio Fadli Ramadhan[1], Rakhmat Kurniawan[2]
[1,2]Computer Science Study Program, Faculty of Science and Technology, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

A R T I C L E I N F O

Article history:

Received Jul 23, 2025
Revised Aug 11, 2025
Accepted Aug 19, 2025

Keywords:

Analysis;
Data Privacy;
Naïve Bayes;
Sentiment;
Social Media;
TF-ID.

ABSTRACT

The development of information technology has triggered public concern about data privacy issues, especially on social media such as X (formerly Twitter). The rampant leaks of personal data have driven the need for a deeper understanding of public opinion. This study aims to analyze public sentiment towards data privacy issues by applying the Naïve Bayes algorithm. The formulation of the problem includes how the public perceives data privacy, how the algorithm performs in classifying sentiment, and how the evaluation results of the model used are. This study uses a quantitative method with a text mining and machine learning approach. Data were taken through crawling techniques on 1,500 tweets related to data privacy. The pre-processing stages were carried out through cleaning, tokenizing, normalization, stopword removal, and stemming. Furthermore, the data was labeled using the InsetLexicon dictionary and weighted using the TF-IDF method. The classification model was built using the Naïve Bayes algorithm and evaluated using accuracy, precision, recall, and f1-score metrics. The results showed that the majority of public opinion on data privacy issues was negative, reflecting concerns over the weak protection of personal data. The Naïve Bayes model performed quite well in sentiment classification. This research is useful in providing insight to the government and digital service providers in developing data protection policies that are more responsive to public opinion.

Corresponding Author:

Rio Fadli Ramadhan,
Computer Science Study Program,
Faculty of Science and Technology,
Universitas Islam Negeri Sumatera Utara,
Jl. Wipe. Golf No.120, Kp. Tengah, kec. Pancur Batu, Kab. Deli Serdang, Sumatera Utara 20353, Indonesia
Email: riofadly87@gmail.com

1. INTRODUCTION

The increasing number of data breaches in Indonesia reflects the weakness of information security systems and has sparked growing public concern, especially on social media. This situation creates an urgency to develop Naïve Bayes-based sentiment analysis methods to better understand public perceptions, opinions, and levels of trust regarding how the government and digital service providers handle data breach issues. By automatically and accurately analyzing public sentiment, the research findings can serve as a foundation for stakeholders to formulate more effective communication strategies and policies to address public concerns and enhance trust in data security in Indonesia.

The rapid development of information and communication technology has fundamentally changed the way people interact, work, and access information. In today's digital era, nearly every aspect of life is integrated with digital technology, from financial transactions and online education to digital healthcare services to daily social activities via social media (Wulan Sari et al., 2023). This

convenience is supported by large-scale data storage and processing, which requires public trust in the systems that manage their personal information (Fuad Amirullah et al., 2023).

However, with increasing digitalization, data privacy issues have become a serious challenge for both individuals and institutions (Wibawa et al., 2019). Data privacy refers to an individual's right to control how their personal information is collected, stored, used, and disseminated (Ramdan et al., 2023). In Indonesia, concerns about personal data misuse are increasing due to the rise in data leaks, such as hacking of public service applications, misuse of customer data by digital service providers, and weak oversight of data processing by third parties (Firdaus et al., 2023).

Based on a report from the Indonesian Cyber Research Institute (2023), more than 60% of internet users in Indonesia stated that they had experienced or felt threatened by data leaks.(Diana Dwi Rahayu et al., 2024) Cases such as the e-KTP data leak, BPJS Kesehatan customer data leak, and data leaks from online loan applications demonstrate that data security remains a significant challenge for the government and digital service providers (Maharani & Triayudi, 2022). This not only results in material losses but also undermines public trust in the use of technology (Hairunnisa & Syaka, 2022).

Previous research employing the Naïve Bayes algorithm for sentiment analysis has generally focused on broader topics such as consumer reviews, political discourse, or general public opinion, often demonstrating the algorithm’s reliability in handling large-scale text data with relatively simple computational requirements. However, studies specifically applying Naïve Bayes to the context of data privacy and security issues, particularly in Indonesia, remain limited. This research positions its contribution by not only reaffirming the effectiveness of Naïve Bayes in classifying sentiments but also by contextualizing its application to the urgent and highly sensitive issue of personal data protection. In doing so, it fills an important gap by providing empirical evidence of public concerns in the Indonesian digital landscape, thereby extending the scope of sentiment analysis research into a domain of significant societal and policy relevance.

In response to these threats, the Indonesian government passed Law No. 27 of 2022 concerning Personal Data Protection (PDP Law). This law is considered a significant milestone in upholding citizens' privacy rights in the digital realm (Azizah et al., 2019). However, the implementation and effectiveness of the PDP Law still pose many questions, particularly regarding law enforcement, public awareness, and the readiness of relevant institutions to manage information systems that comply with data protection regulations (Zufria et al., 2024).

On the other hand, social media has become a primary channel for people to express their opinions and experiences regarding data privacy issues (Kosasih & Alberto, 2021). Platforms like X (formerly Twitter), Facebook, and Instagram have become public discussion spaces where various views, criticisms, and concerns about personal data security are openly expressed (Christopher, 2025). Through social media, people respond to various issues such as overly invasive app terms and conditions, ads that appear based on user behavior, and data mining practices by large technology companies (Khotimah & Ula, 2023).

In the context of data privacy issues, the Quran does not explicitly mention it because the concept of digital was not yet present at the time of revelation (Setiawan et al., 2024). However, the fundamental values of trustworthiness, justice, and responsibility for information are highly relevant as a basis for ethical personal data protection (Anugerah et al., 2024). One of the Quranic verses most often cited as a basis for ethical information management and the responsibility to safeguard trusts, including personal data, is Surah An-Nisa, verse 58.

إِنَّ اللَّهَ يَأْمُرُكُمْ أَنْ تُؤَدُّوا الْأَمَانَاتِ إِلَىٰ أَهْلِهَا ۙ وَإِذَا حَكَمْتُم بَيْنَ النَّاسِ أَنْ تَحْكُمُوا بِالْعَدْلِ ۚ إِنَّ اللَّهَ نِعِمَّا يَعِظُكُم بِهِ ۗ إِنَّ اللَّهَ كَانَ سَمِيعًا بَصِيرًا ﴿٥٨﴾

Meaning: "Indeed, Allah commands you to convey trust to its owner. When you determine a law between people, you must determine it fairly. Indeed, Allah gives you the best teaching. Indeed, Allah is All-Hearing, All-Seeing."

This verse emphasizes the importance of maintaining trust and upholding justice, which is relevant to data privacy issues in the digital age (Rachman & Handayani, 2021). In this context, personal data provided by users to digital platforms constitutes a trust that must be safeguarded and not misused. Islam commands that every trust be returned to its rightful owner and prohibits

any form of betrayal of trust (Titimeidara & Hadikurniawati, 2021). Furthermore, this verse emphasizes the need for justice in making decisions, including in the management and use of data, so that authorities do not misuse information for unilateral gain (Putro et al., 2020). Thus, the principles contained in this verse serve as an ethical foundation for personal data protection and demand moral responsibility from data managers to consistently maintain public trust in a fair and trustworthy manner.

The high level of public activity on social media generates a large amount of opinion data with significant potential for analysis. However, this data is unstructured, scattered, and diverse in context and meaning (Permana et al., 2024). Therefore, a manual approach to understanding public opinion is highly inefficient and prone to biased interpretation. To address this issue, automated, technology-based approaches, such as text mining and sentiment analysis, are needed to extract and classify public opinion accurately and systematically (Ihsan et al., 2024).

Sentiment analysis is a technique used to identify a person's attitudes, feelings, or opinions toward a particular topic, product, or issue. In this context, sentiment analysis of data privacy issues can provide a general overview of how the public views policies, practices, and events related to data protection in the digital world (Asfi & Fitrianingsih, 2020). This approach can help governments, non-profit organizations, and the private sector understand public perceptions and formulate more targeted communication strategies or policies.

Previous research related to sentiment analysis of campaigns and also serves as a reference for the current research is research by Puad et al. (2023) entitled "Analysis of Public Sentiment on Twitter Regarding the 2024 General Election Using the Naïve Bayes Algorithm." In this study, researchers obtained comment data related to the 2024 general election using a crawling technique with a total of 6000 data. Then the researchers pre-processed the data first such as cleaning, case folding, tokenizing, stopword removal and stemming (Alfarezy et al., 2023). Then the data was processed using the Naïve Bayes algorithm and then weighted words with the TF-IDF algorithm. Then obtained 331 positive data labels, 261 negative labels, and 825 neutral labels, the results showed that the model with 90:10 data had the highest accuracy (Puad et al., 2023).

In this study, the author will adopt a similar method to analyze public sentiment on data privacy issues. The data used are tweets with relevant keywords, which will be classified into three sentiment categories: positive, negative, and neutral, using the Naïve Bayes algorithm. Naïve Bayes is a method used to classify a set of data or documents (Syarli & Muin, 2020). This algorithm utilizes probability and statistics methods proposed by British scientist Thomas Bayes. This algorithm utilizes probability and statistics to analyze sentiment patterns in the data, which can provide an overview of the level of public trust and expectations towards the government (Arya & Zufria, 2024).

This study aims to apply the Naïve Bayes algorithm in analyzing public sentiment towards data privacy issues on social media X. By using crawled data from public conversations on the platform, the author will carry out data pre-processing processes such as tokenization, stemming, and weighting using TF-IDF. The classification results will be evaluated using accuracy, precision, recall, and f1-score metrics to assess the model's performance in automatically classifying public opinion.

This research is expected to provide deeper insights into public perceptions of data privacy and contribute to efforts to improve digital literacy and data protection policies in Indonesia. Technology-based approaches such as sentiment analysis can help governments and organizations better respond to public aspirations and concerns as they navigate the increasingly complex privacy challenges of the digital age.

2. RESEARCH METHOD

This research uses a quantitative approach, emphasizing the collection and analysis of numerical data to identify statistical and causal relationships between variables. This approach was chosen because it can provide measurable and objective results, especially when processing large amounts of data. Quantitative research also allows researchers to systematically test hypotheses using statistical models, so that the results can be used as a basis for data-driven decision-making (Hartono, 2019). This study specifically utilizes the Naïve Bayes Classifier algorithm to analyze text data to determine public sentiment regarding personal data privacy issues in the digital era.

The first stage of this research begins with problem identification, which focuses on the increasing public concern about the issue of personal data leakage due to the use of digital services such as social media, public applications, and e-commerce platforms. This issue has become a major concern in the digital public space, especially on social media platform X (formerly Twitter). After the problem was identified, data collection was carried out using web scraping techniques using Google Colab. The data collected consisted of 1,500 tweets containing keywords such as "data privacy," "data leakage," and "PDP Law," which were relevant to the research topic.

The collected data then undergoes text pre-processing, including cleaning, case folding, stopword removal, tokenizing, and stemming to transform the raw text into clean, processable data. Afterward, a word weighting process is performed using the Term Frequency–Inverse Document Frequency (TF-IDF) method to weight words that have a significant influence on distinguishing sentiment. The processed data is then fed into the Naïve Bayes algorithm to be classified into three sentiment categories: positive, negative, and neutral. This algorithm was chosen for its simplicity, processing speed, and effectiveness in classifying large and complex text data.

Next, the model's performance was tested using test data. The evaluation was conducted through measurements of accuracy, precision, recall, and f1-score to determine the model's ability to correctly recognize sentiment in text. The classification results will be further analyzed to determine public perceptions of personal data protection, the effectiveness of government policies, and the level of public trust in digital service providers. With a structured quantitative approach and the support of machine learning-based classification techniques, this research is expected to provide a comprehensive picture of public opinion on crucial issues in the digital era.

3. RESULTS AND DISCUSSIONS

This study successfully collected 1,500 tweets from social media platform X (formerly Twitter) containing keywords related to data privacy issues. After web scraping, the raw data underwent a cleaning process and removed irrelevant columns such as created_at and favorite_count, leaving only the full_text column as the primary source of sentiment analysis. Initial observations of tweet content indicate that data privacy issues have received diverse responses from the public, ranging from complaints about personal data leaks, concerns about information misuse, to aspirations for stronger legal protection. This indicates that digital privacy is a significant issue in today's public opinion.

After the data was cleaned and prepared, preprocessing steps were performed, including cleaning, case folding, tokenizing, normalization, stopword removal, and stemming to ensure the data was in optimal form before classification. The sentiment labeling process was performed using the InsetLexicon dictionary, which automatically labels each tweet as "positive" or "negative" based on the keywords it contains. A model was then built using the Naïve Bayes Classifier algorithm, which calculates the probability of a text belonging to a particular class based on the distribution of words in the training data. For word weighting, the TF-IDF method was used to strengthen the influence of important words that appear uniquely in each text document.

Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. Test results showed that the Naïve Bayes algorithm was able to classify sentiment data with a fairly high level of accuracy, indicating that the model performed optimally in capturing public opinion trends regarding data privacy issues. Overall, the data distribution indicated that negative sentiment dominated, reflecting public concerns about the lack of personal data protection in the digital age. On the other hand, a small proportion of tweets contained positive sentiment, generally responding to company or government initiatives deemed proactive in maintaining user data security.

Figure 1. Visualization of sentiment analysis count

After all comment data has gone through a preprocessing process that includes text cleaning, tokenization, normalization, stopword removal, stemming, and detokenization, the next step is sentiment labeling using the InSetLexicon dictionary. This labeling process is performed automatically by calculating the total polarity weights of the words in each comment. If the sum of the word weights is positive, the comment is categorized as positive sentiment; conversely, if the value is negative, the comment is labeled as negative sentiment.

Based on the labeling results, a visualization of the distribution of comments based on their sentiment categories was performed. This visualization process was created using the Python programming language through the Google Colaboratory platform. The results are presented in the form of a bar chart in Figure 1, which shows the proportion of positive and negative comments. The visualization shows that the number of comments with positive sentiment outnumbers the number of negative comments. This finding indicates that the majority of users' responses tended to be supportive, informative, or expressed appreciation in the context of data privacy issues.

However, negative comments remain a significant proportion and cannot be ignored. The presence of negative comments often reflects complaints, criticism, or unpleasant experiences experienced by users. Therefore, the results of this sentiment analysis not only provide a general overview of public perception but can also serve as an important foundation for relevant parties to improve services, enhance communication, and design more transparent and user-friendly privacy policies.

After the sentiment labeling process was carried out using the InSetLexicon dictionary, the next step was to visualize the proportion of comments that fell into the positive and negative sentiment categories. This visualization aims to provide an overview of the trends in public opinion or perception contained in the analyzed comment data. Based on the labeling results, each comment was assigned a polarity value and classified into two main classes: "positive" and "negative." This entire process was automated using Python programming on the Google Colaboratory platform.

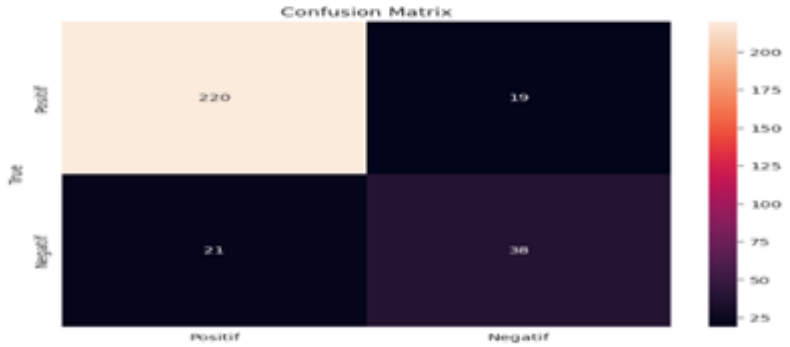

Figure 2. Positive wordcloud

Figure 3. Negative wordcloud

After sentiment testing using the Naive Bayes algorithm, the next step is to conduct an evaluation using a confusion matrix to calculate accuracy, precision, recall, and F1 score by comparing the resulting classification labels with the actual labels in the sentiment dataset related to data privacy issues in X, namely:

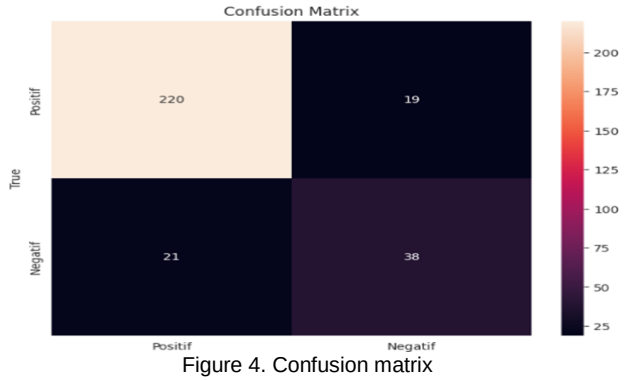

Figure 4. Confusion matrix

In Figure 4, the classification results for finding accuracy, precision, recall and F1-score follow the following formula:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{220 + 38}{220 + 38 + 19 + 21} = 0.8658$$

$$precision = \frac{TP}{TP + FP} = \frac{220}{220 + 21} = 0.91286$$

$$recall = \frac{TP}{TP + FN} = \frac{220}{220 + 19} = 0.92050$$

$$F1 - Score = 2\ x\ \frac{presicion\ x\ recall}{precision + recall} = 2\ x\ \frac{0.91286\ x\ 0.92050}{0.91286 + 0.92050} = 0.9159$$

The comment accuracy for the dataset is shown in the image below where it is 0.8657, precision is 0.91, recall is 0.92 and F1-Score is 0.92.

Hasil Akurasi: 0.8657718120805369

	precision	recall	f1-score	support
Negatif	0.91	0.92	0.92	239
Positif	0.67	0.64	0.66	59
accuracy			0.87	298
macro avg	0.79	0.78	0.79	298
weighted avg	0.86	0.87	0.86	298

Figure 5. Confusion matrix results

The confusion matrix results in Figure 5 show the performance of the sentiment classification model on data privacy issues using the Naïve Bayes algorithm. Of the 298 test data sets, the model successfully classified 220 positive data sets correctly and 38 negative data sets accurately, with an overall accuracy rate of 86.57%. In the Negative class, the model performed very well with a precision of 0.91 and a recall of 0.92, resulting in a consistent f1-score of 0.92. Meanwhile, in the Positive class, the precision was only 0.67 and the recall was 0.64, indicating that the model had more difficulty recognizing comments with positive sentiment, which was also reflected in the f1-score of 0.66. The macro average f1-score of 0.79 indicates fairly stable model performance across classes, although there were performance disparities between categories. The weighted average f1-score of 0.86 indicates that the largest contribution to model performance came from the Negative class, which indeed had a larger amount of data. This shows that the model has a tendency to predict negative sentiment more accurately than positive sentiment on data privacy issues.

Overall, this study successfully demonstrated that the Naïve Bayes algorithm can be used effectively to classify sentiments on comments related to data privacy issues on platform X. The text preprocessing process, which includes cleaning, tokenizing, normalization, stopword removal, stemming, and detokenization, proved crucial in creating clean and consistent data before sentiment analysis. The evaluation results showed that the model performed well in identifying negative comments, although improvements in recognizing positive sentiments are still needed to achieve a more balanced classification. Therefore, this approach can be a starting point for developing a more intelligent and responsive public opinion monitoring system for sensitive issues such as data privacy on social media.

4. CONCLUSION

Based on the research results, the application of the Naïve Bayes algorithm in sentiment analysis on data privacy issues on social media X (Twitter) successfully provided an overview of public opinion trends. From 1,486 cleaned and processed data, it was found that most comments tended to be negative, indicating high public concern about the protection of their personal data. The built model achieved an accuracy of 86.57% with optimal performance in classifying negative comments, although it still showed weaknesses in identifying positive ones. This highlights the need to strengthen the model's ability to recognize diverse sentiment expressions for a more comprehensive analysis. As a follow-up, further development can be carried out by expanding the scope of data sources, adopting more sophisticated classification algorithms, and applying advanced text processing techniques such as word embedding. Addressing the imbalance between positive and negative comments through balancing techniques is also crucial to reduce bias, alongside the use of cross-validation and error analysis for deeper performance evaluation.

From a practical perspective, these findings have important implications for policymakers. By leveraging sentiment analysis as a decision-support tool, policymakers can better capture and understand public concerns, enabling them to formulate data protection regulations that are not only technically robust but also responsive to the voices of citizens. In this way, regulatory frameworks can be aligned more closely with public expectations, thereby strengthening trust and legitimacy in data governance within the digital age.

REFERENCES

Alfarezy, R., Ermatita, E., & Wadu, R. M. B. (2023). Implementasi Algoritma Naïve Bayes Untuk Analisis Klasifikasi Survei Kesehatan Mental (Studi Kasus: Open Sourcing Mental Illness). *Informatik : Jurnal Ilmu Komputer*, *19*(1), 1–10. https://doi.org/10.52958/iftk.v19i1.4696

Anugerah, S. M., Wijaya, R., & Bijaksana, M. A. (2024). Sentimen Analysis Social Media for Disaster using Naïve Bayes and IndoBERT. *INTEK: Jurnal Penelitian*, *11*(1), 51–58. https://doi.org/10.31963/intek.v11i1.4771

Arya, D., & Zufria, I. (2024). *Analisis Sentimen Terhadap Program Kampanye Desak Anies Di X Menggunakan Naïve Bayes*. *5*(1), 104–115. https://doi.org/10.30865/klik.v5i1.2085

Asfi, M., & Fitrianingsih, N. (2020). InfoTekJar : Jurnal Nasional Informatika dan Implementasi Algoritma Naive Bayes Classifier sebagai Sistem Rekomendasi Pembimbing Skripsi. *InfoTekJar*, *5*.

Azizah, N., Goejantoro, R., & Sifriyani. (2019). Metode Naive Bayes Dengan PendekatanDistribusi Gauss Untuk Klasifikasi PeminatanPeserta Didik. *Prosiding Seminar Nasional Matematika Dan Statistika*, *1*, 1–7. https://jurnal.fmipa.unmul.ac.id/index.php/SNMSA/article/view/520/217

Christopher, B. (2025). Analisi Komparasi Metode K-Nearest Neighbor dan Naïve Bayes Classifier Berbasis Optimasi Randomized Search Dalam Klasifikasi Berita Hoaks. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, *12*(2), 272–283. https://doi.org/10.35957/jatisi.v12i2.11371

Diana Dwi Rahayu, Muhammad Fatchan, & Alfonsus Ligouri. (2024). Analisis Sentimen Twitter Terpilihnya Prabowo - Gibran Menggunakan Metode Neural Network. *Tematik*, *11*(1), 85–91. https://doi.org/10.38204/tematik.v11i1.1943

Firdaus, A. A., Yudhana, A., & Riadi, I. (2023). Sentiment Analysis on 2024 Presidential Election Projection using Support Vector Machine Method. *Decode: Jurnal Pendidikan Teknologi Informasi*, *3*(2), 236–245.

Fuad Amirullah, Syariful Alam, & M.Imam Sulistyo S. (2023). Analisis Sentimen Terhadap Kinerja KPU Menjelang Pemilu 2024 Berdasarkan Opini Twitter Menggunakan Naïve Bayes. *STORAGE: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, *2*(3), 69–76. https://doi.org/10.55123/storage.v2i3.2293

Hairunnisa, H., & Syaka, W. A. (2022). Analisis Komunikasi Politik Dalam Percepatan Pembangunan Ibu Kota Nusantara (IKN) Menuju Kota Berkelanjutan. *Journal of Government and Politics (JGOP)*, *4*(1), 1. https://doi.org/10.31764/jgop.v4i1.8193

Ihsan, T., Ariwibowo, M. F., Indra, A. M., & ... (2024). Analisis Deskripsi Pemilihan Tema Konten untuk Pemasaran Sosial Media Instagram dalam peningkatan personal Branding. *MDP Student …*, 924–929.

Khotimah, K., & Ula, D. M. (2023). Triwikrama: Jurnal Ilmu Sosial. *Triwikrama: Jurnal Ilmu Sosial*, *01*(11), 40–50.

Kosasih, R., & Alberto, A. (2021). Sentiment analysis of game product on shopee using the TF-IDF method and naive bayes classifier. *ILKOM Jurnal Ilmiah*, *13*(2), 101–109. https://doi.org/10.33096/ilkom.v13i2.721.101-109

Maharani, A. P., & Triayudi, A. (2022). Sentiment Analysis of Indonesian Digital Payment Customer Satisfaction Towards GOPAY, DANA, and ShopeePay Using Naïve Bayes and K-Nearest Neighbour Methods. *Jurnal Media Informatika Budidarma*, *6*(1), 672. https://doi.org/10.30865/mib.v6i1.3545

Permana, A. A., Taufiq, R., Destriana, R., & Nur'aini, A. (2024). Implementasi Algortima Naïve Bayes Untuk Prediksi Kelulusan Mahasiswa. *Jurnal Teknik*, *13*(1), 65–70. https://jurnal.umt.ac.id/index.php/jt/article/view/10996

Puad, S., Garno, G., & Susilo Yuda Irawan, A. (2023). Analisis Sentimen Masyarakat Pada Twitter Terhadap Pemilihan Umum 2024 Menggunakan Algoritma Naïve Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(3), 1560–1566. https://doi.org/10.36040/jati.v7i3.6920

Putro, H. F., Vulandari, R. T., & Saptomo, W. L. Y. (2020). Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan. *Jurnal Teknologi Informasi Dan Komunikasi (TIKomSiN)*, *8*(2). https://doi.org/10.30646/tikomsin.v8i2.500

Rachman, R., & Handayani, R. N. (2021). Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM. *Jurnal Informatika*, *8*(2), 111–122. https://doi.org/10.31294/ji.v8i2.10494

Ramdan, A. S., Fadzilah, A. F., Haleel, A., & Misbah, E. (2023). Analisis Perbandingan Komunikasi Politik Ganjar Pranowo, Prabowo Subianto, Anies Baswedan Melalui Video "3 Bacapres Bicara Gagasan" di Channel YouTube Najwa Shihab. *Prosiding Seminar Nasional*, 780–791.

Setiawan, I. B., Maulindar, J., & Nurchim, N. (2024). Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Pada Aplikasi Kesehatan Digital. *G-Tech: Jurnal Teknologi Terapan*, *8*(4), 2301–2312. https://doi.org/10.70609/gtech.v8i4.5020

Syarli, & Muin, A. A. (2020). Metode Naive Bayes Untuk Prediksi Kelulusan. *Jurnal Ilmiah Ilmu Komputer*, *2*(1), 22–26. https://media.neliti.com/media/publications/283828-metode-naive-bayes-untuk-prediksi-kelulu-139fcfea.pdf

Titimeidara, M. Y., & Hadikurniawati, W. (2021). Implementasi Metode Naïve Bayes Classifier Untuk Klasifikasi Status Gizi Stunting Pada Balita. *Jurnal Ilmiah Informatika*, *9*(01), 54–59. https://doi.org/10.33884/jif.v9i01.3741

Wibawa, A. P., Kurniawan, A. C., Murti, D. M. P., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., &

Nugraha, Y. R. (2019). Naïve Bayes Classifier for Journal Quartile Classification. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, *7*(2), 91. https://doi.org/10.3991/ijes.v7i2.10659

Wulan Sari, D., Sari Putri, M., & Nurlaili, N. (2023). Relevansi Pendidikan Islam Di Era Digital Dalam Menavigasi Tantangan Modern. *Science and Education Journal (SICEDU)*, *2*(2), 372–380. https://doi.org/10.31004/sicedu.v2i2.129

Zufria, I., Lubis, A. H., Febiyaula, S. S., Islam, U., & Sumatera, N. (2024). *Kepolisian Republik Indonesia Menggunakan*. *4307*(August), 1266–1272.