Assessing Academic Information System Performance Through Sentiment Analysis

By Zafira Thuraya, Ali Ibrahim, Yadi Utama, Dwi Rosa Indah

Assessing Academic Information System Performance Through Sentiment Analysis

[1]Zafira Thuraya, [2]Ali Ibrahim*,[3]Yadi Utama, [4]Dwi Rosa Indah
[1,3,4] Information System, Faculty of Computer Science, Sriwijaya University, Indonesia
[2]Master of Computer Science, Faculty of Computer Science, Sriwijaya University, Indonesia
[1,2,3,4] Jalan Palembang-Prabumulih, KM 32 Inderalaya, Kab. Ogan Ilir, Sumatera Selatan, Indonesia
*e-mail: aliibrahim@unsri.ac.id

(***received***: ?, ***revised***: ?, ***accepted***: ? diisi oleh editor)

Abstrak

Sistem Informasi Akademik (SIMAK) di Universitas Sriwijaya memiliki peranan yang signifikan dalam mendukung kegiatan akademik mahasiswa, namun masih terdapat beberapa kendala teknis yang berdampak pada kepuasan pengguna. Penelitian ini bertujuan untuk menganalisis sentimen mahasiswa terhadap SIMAK dengan menggunakan metode Naïve Bayes. Data diperoleh dari Twitter melalui teknik web scraping dan selanjutnya diproses dengan metode pembersihan teks serta pembobotan TF-IDF. Hasil analisis menunjukkan bahwa model yang digunakan mencapai akurasi sebesar 65%, dengan precision untuk sentimen negatif sebesar 63% dan recall mencapai 100%. Sementara itu, precision untuk sentimen positif hanya 12,5% dengan F1-score sebesar 22,2%. Model ini lebih efektif dalam mendeteksi keluhan pengguna yang berkaitan dengan gangguan server, keterlambatan data, serta kesulitan dalam pengisian KRS dan akses nilai. Oleh karena itu, perbaikan teknis sangat diperlukan untuk meningkatkan stabilitas sistem. Penelitian selanjutnya disarankan untuk mengeksplorasi metode klasifikasi alternatif seperti Support Vector Machine (SVM) guna meningkatkan akurasi dalam analisis sentimen.

Kata kunci: Sistem Informasi, Anlisis sentimen, Naive Bayes, Sosial Media, Twitter

Abstract

The Academic Information System (SIMAK) at Sriwijaya University plays a crucial role in supporting students' academic activities; however, it continues to encounter technical challenges that impact user satisfaction. This study analyzes student sentiment towards SIMAK utilizing the Naïve Bayes method. Data was collected from Twitter through web scraping and processed using text cleaning techniques and TF-IDF weighting. The results indicate that the model achieved an accuracy of 65%, with a negative sentiment precision of 63% and a recall of 100%, while the positive sentiment precision was only 12.5%, resulting in an F1-score of 22.2%. The model proved to be more effective in identifying user complaints related to server disruptions, data delays, and difficulties in filling out course registration forms and accessing grades. Technical improvements are necessary to enhance system stability. Future research is recommended to explore alternative classification methods, such as Support Vector Machine (SVM), to improve sentiment analysis accuracy.

Keywords: *Information System, Sentiment analysis, Naïve Bayes, Social Media, Twitter*

1 Introduction

According to data from Reportal, internet usage in Indonesia has reached 185.3 million individuals, reflecting an increase of approximately 0.8% or 1.5 million people over the past year [1]. The integration of the internet within educational institutions presents significant opportunities for the advancement of digital learning in Indonesia, including innovations in both the learning and evaluation processes [2]. One notable application of the internet in higher education is the Academic Information System (SIMAK), which facilitates access to information and communication among students, faculty, and administrative staff [3].

Universitas Sriwijaya has developed SIMAK through its internal IT team to enhance the management of student academic data [4]. This system serves various academic purposes, such as filling out the Study Plan Card (KRS), accessing grades on the Study Result Card (KHS), and reviewing

ISSN:2302-8149
e-ISSN:2540-9719

academic transcripts [5]. However, in practice, several challenges persist, including incomplete information, a less intuitive interface, difficult navigation, delays in data updates, and technical disruptions that hinder the user experience [6]. These issues may affect student satisfaction with SIMAK, ultimately impacting the overall effectiveness of academic services. Therefore, it is essential to conduct evaluations based on student feedback to gain insights into their perceptions of this system.

To assess the extent to which SIMAK meets the needs of its users, an analysis of the services provided is essential for system improvement. In the digital age, social media has emerged as a primary platform for students to express their opinions and experiences, particularly regarding academic services [7]. One frequently utilized platform is X (formerly known as Twitter), which allows users to publish tweets containing text, images, videos, or links [8]. With its extensive user base and role in public discourse, X serves as a relevant data source for various analyses, including sentiment analysis related to academic services in higher education [9].

Sentiment analysis is a widely employed method for examining textual data, such as comments, reviews, tweets, forums, and articles [10]. In this study, the Naïve Bayes method is utilized as the primary approach to analyze student opinions regarding SIMAK. This algorithm operates on the principle of probability, where the Naïve Bayes model identifies the highest probability values to accurately categorize test data [11].

Previous research has demonstrated that the Naïve Bayes method has been successfully applied in various sentiment analysis contexts. For instance, a study examining public opinion on Twitter regarding the childfree issue reported a Naïve Bayes model accuracy of 69.69% [12]. Another study compared this method with the C4.5 algorithm in analyzing sentiment from users of the JKN mobile application, where Naïve Bayes achieved an accuracy of 79.95%, surpassing C4.5, which recorded 79.17% [13]. However, these studies primarily focused on social issues and healthcare services, while sentiment analysis related to academic information systems, particularly at Universitas Sriwijaya, remains underexplored.

This research introduces a new perspective by highlighting the experiences of Universitas Sriwijaya students in utilizing SIMAK, an area that has not been extensively studied in prior research. Consequently, conducting sentiment analysis on student evaluations of SIMAK using the Naïve Bayes method is crucial for understanding user perceptions. The findings of this study are anticipated to provide valuable insights for academic information system developers in designing improvements that better align with student needs. Furthermore, this research may serve as a reference for other educational institutions in developing more responsive, adaptive, and data-driven academic information systems to support innovative academic services in line with technological advancements and user expectations.

2 Literature Review

Research conducted by Tri et al. utilized a questionnaire as a source of sentiment data. The objective of this study was to classify student opinions regarding instructors based on various factors, including mastery of the subject matter, explanation skills, teaching methods, punctuality, and other relevant aspects. The Naïve Bayes Classifier was employed in this research to categorize student sentiments into three distinct groups: positive, negative, and neutral. The outcome of this study is a web-based system that presents a summary of the instructor questionnaires, accessible at any time and from any location [14]

Another study was carried out by focusing on the Academic Information System at the Garut Institute of Technology [15] This research aimed to assess user sentiment towards the SIAM application to assist developers in enhancing the system. However, the research methodology was not elaborated upon in detail, and the study primarily highlighted the tools utilized. The findings indicated that positive sentiment accounted for 57.14%, negative sentiment for 37.14%, and neutral sentiment for 5.71%, which were visualized in the form of a word cloud.

Additionally, Dualu et al. investigated the sentiment classification of students regarding the Academic Information System (SIAKAD) at STIMIKOM Stella Maris Sumba. This system is accessible to students, staff, and faculty for academic purposes. Data were collected through a questionnaire, yielding 112 comments, of which 62 expressed positive sentiment. This research employed the Naïve Bayes algorithm with RapidMiner as the software tool, and the testing results demonstrated an accuracy rate of 95.38% [16]

ISSN:2302-8149
e-ISSN:2540-9719

Furthermore, Luqma et al. conducted a study on the SIAM at Brawijaya University using the K-Nearest Neighbor algorithm with feature selection via the Chi-square method. The research dataset was sourced from Twitter, containing complaints regarding service disruptions of SIAM UB. The classification results revealed an accuracy level of 86% [17]

In contrast to previous research, the study conducted by Yoga et al. centers on the analysis of student reviews regarding faculty performance in both mandatory and elective courses. The research data was collected through a survey conducted at SIAM UB, yielding a total of 3,805 reviews. The Random Forest algorithm was employed in this study, achieving an accuracy rate of 90%, precision of 99%, recall of 96%, and an F1-Score of 97% [18].

From the literature review previously discussed, it can be concluded that the Naïve Bayes method is widely utilized in sentiment analysis of student feedback on academic information systems. The accuracy results vary depending on additional methods such as cross-validation or feature selection. These studies contribute to a better understanding of student perceptions regarding academic services and serve as a foundation for the development of more accurate sentiment analysis models in the future.

3 Research Method

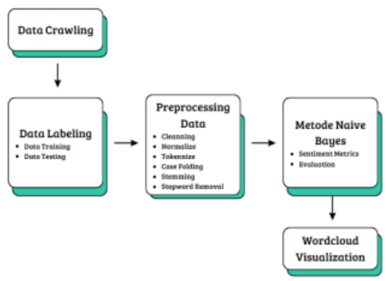

Figure 1 Research Metodelogy

The stages of this research can be described as in Figure 1 flowchart of the research stages. The research stages include explaining each stage in the research process starting from dataset collection to wordcloud visualization.

3.1 Data Crawling

The study implements the Naive Bayes Classifier method utilizing Google Colab, based on the collected student comment review data. The Naive Bayes algorithm will predict outcomes by leveraging the initial probabilities of each label derived from the training set, along with the contribution of each feature [19]. The dataset employed in this research consists of student reviews regarding the Academic Information System of Sriwijaya University, gathered from social media platform Twitter through the use of an Application Programming Interface (API) [20].

3.2 Labelling Data

The collected data is categorized into two segments, with 70% designated as labeled training data and 30% as unlabeled testing data, a method known as the hold-out technique [21]. The training data is manually labeled with positive and negative sentiments, followed by a pre-processing phase. Due to the irregular structure of text data, several steps are necessary to prepare the text for a more structured format [22]. This division aims to enable the model to learn patterns from the training data and then be tested on unseen data to evaluate its performance [23].

3.3 Pre – processing Data

The data pre-processing stage aims to ensure that the data to be used in the analytical model or machine learning is clean, relevant, and in the right format. Cleansing data use eliminating irrelevant components, empty values, or noise, including emoticons, hashtags, usernames, and URLs [24]. Normalization seeks to enhance consistency in text analysis by converting misspelled words, abbreviations, or slang into their standard forms, thereby ensuring that the analysis results are more valid and reliable [25]

Tokenizing uses spaces to divide the comment text into separate words [26] Case folding involves converting all characters to lowercase to avoid confusion and interference with the results [27]. Stemming reduces words to their root forms to improve the consistency of the analysis [28]. Stopword removal entails eliminating frequently occurring words that do not contribute to the analysis, such as 'di', 'yang', 'ke', 'dari', and others [29]

3.4 Naïve Bayes Method

Based on user sentiment, which is converted into a numerical format using Term Frequency-Inverse Document Frequency (TF-IDF) [30] . This allows the algorithm to evaluate the significance of each word. The simplicity, computational efficiency, and resilience of Naive Bayes Classifier (NBC) against irrelevant features and missing data ensure its optimal performance, even in resource-constrained environments [31]. In this study, a multinational approach to Naive Bayes is employed. The equation for NBC in sentiment classification can be expressed as follows.

$$P(H|X) = \frac{P(H|X)\,P(H)}{P(X)} \quad (1)$$

3.5 Confusion Matrix

To evaluate the model, this matrix comprises four variables or components that is TP (True Positive) and TN (True Negative) variables indicate the correct predictions made by the model overall, while the FP (False Positive) and FN variables represent the total number of incorrect predictions generated by the model [32]. Once the values in the confusion matrix are established, the next step is to assess the analysis results by calculating recall, precision, accuracy, and F-measure[33].

Table 1 Confusion Matrix

	Predict Positive	Predict Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	False Negative (FN)

3.6 Wordcloud

Word cloud like visual representation of frequently occurring or dominant words associated with a theme[34].Word cloud visualization is expected to provide better insight and understanding of the interpretation of a particular word or theme.

4 Results and Analysis

This research utilizes data sourced from Twitter by leveraging the Twitter API, which facilitates automated data collection. The data collection period spans from March 2023 to August 2024, focusing on the keyword "simak unsri," resulting in a total of 104 sentiment instances. Following the data crawling phase, manual labeling was conducted to prepare the data for training, categorizing the sentiments into two classes: positive and negative. The results are presented in Table 2 below.

Table 2 Labelling Data

Label	Total
Positive	26
Negative	37
Total Training	63

Unlabeled data will be set to null to test the accuracy of the Naïve Bayes model. The training data is labeled prior to the pre-processing stage to prevent sentiment confusion, as some words may be removed during the stopword phase. The text pre-processing stage includes several steps, such as cleaning, standard word normalization, tokenization, case folding, stemming, and stopword removal. After the cleaning process, the resulting clean data, which consists of 91 sentiments.

Table 3 Example Preprocessing

Preprocessing	Result

Cleaning	Simak unsri kenapa terjadi kesalahan server terus tiap klik krs
Normalize	Simak unsri kenapa terjadi kesalahan server terus setiap klik krs
Tokenize	['Simak', 'unsri', 'kenapa', 'terjadi', 'kesalahan', 'server', 'terus', 'setiap', 'klik', 'krs']
Case folding	['simak', 'unsri', 'kenapa', 'terjadi', 'kesalahan', 'server', 'terus', 'setiap', 'klik', 'krs']
Stemming	['simak', 'unsri', 'kenapa', 'jadi', 'salah', 'server', 'terus', 'tiap', 'klik', 'krs']
Stopword	['simak', 'unsri', 'salah', 'server', 'klik', 'krs']

The cleaned data will be processed using the Naïve Bayes modeling approach. The training and testing datasets will be divided using a data split method in Python. Following this division, the vectorization stage will convert the text into a numerical format suitable for processing by the Naïve Bayes model. In this study, the method employed is Term Frequency - Inverse Document Frequency (TF-IDF). This technique not only transforms text into numerical values but also assigns weights to each word based on its frequency in a single document relative to the entire document set. Consequently, words that are more significant for sentiment determination will carry higher weights, while common words that appear frequently across many documents will have lower weights. The output from this vectorization process will serve as the primary input for training the model, which will subsequently be utilized to classify sentiment in the test data.

Once the model has been trained using the training data, the next step is to evaluate its performance with a test dataset that has not been previously utilized. The model will be assessed to determine its ability to recognize the learned patterns and apply them to different texts. The predictions from this model yielded 27 outcomes, with 25 classified as negative sentiment and the remaining 2 as positive sentiment. The results from the Naïve Bayes classifier can be employed to assess the model's effectiveness in classifying user sentiment towards SIMAK, utilizing a confusion matrix. This matrix will provide a visual representation of the model's classification performance, including calculations for accuracy, precision, recall, and F1-score, formatted in a contingency table as shown in Table 4.

Table 4 The performance of connfusion Matrix

	Precision	Recall	F1-Score	Support
Negative	0.63	1.00	0.77	12
Positive	1.00	0.12	0.22	8
Accuracy			0.65	20
Macro avg	0.82	0.56	0.50	20
Weighted avg	0.78	0.65	0.55	20

The table indicates that the model performs well in recognizing the negative class, achieving a recall of 1.00 and an F1-score of 0.77. However, it is less effective in identifying the positive class, as evidenced by a very low recall of 0.12 and an F1-score of 0.22. The overall accuracy of the model stands at 0.65, with a macro average F1-score of 0.50, highlighting an imbalance in the model's performance

ISSN:2302-8149
e-ISSN:2540-9719

across the two classes. This discrepancy may be attributed to data distribution or the model's limitations in addressing the minority class.

To gain further insights into the sentiment within the dataset, data visualization was conducted using a word cloud. The word cloud serves to analyze the frequency of words that frequently appear in user reviews, thereby aiding in the understanding of user preferences and experiences related to the use of SIMAK. By identifying the most relevant keywords, this visualization provides valuable insights for sentiment analysis of the existing data, as illustrated in figure 3.

Figure 3. Wordcloud data

The results indicate that several dominant keywords are related to the use of SIMAK UNSRI, particularly concerning issues of errors or technical difficulties. However, there are also signs that some students have provided positive feedback about the system. As illustrated in the figure 4, words such as "thank you" and "alhamdulillah" stand out, reflecting that many students appreciate SIMAK UNSRI, especially regarding new features and access to grades. These findings suggest that certain elements have successfully met user needs and could represent strengths to be maintained. Nevertheless, there are still indications that the system is not yet fully optimized. Further improvements are necessary to enhance the overall user experience.

Figure 4. Wordcloud label positive

On the other hand, the word cloud representing negative sentiment reviews, as illustrated in the figure 5, highlights the terms that frequently appear in student complaints regarding the system. Words such as "error" and "lag" are prominently sized, indicating the primary issues that are often raised, including system errors, access difficulties, slow loading times, and problems with course registration or grade checking. These complaints underscore the necessity for enhancing system stability, improving access speed, and refining key features to ensure they are more responsive and user-friendly.

Figure 5. Wordcloud label negative

When comparing the word cloud of positive sentiment, it appears that complaints regarding the system remain quite prevalent, both in terms of word count and the intensity of negative language used. Common grievances are related to system performance, such as slow loading times, errors when accessing features, and difficulties in filling out course registration forms and checking grades. Therefore, system improvements are essential to enhance user experience. Efforts should be concentrated on increasing server stability.

5 Conclusion

The findings of this research indicate that the Naïve Bayes model achieved an accuracy of 65% in analyzing student sentiment towards the SIMAK system at Universitas Sriwijaya. This model is more proficient in identifying negative sentiments, but it struggles with accurately detecting positive sentiments, as evidenced by a precision of only 12.5% for positive sentiment, despite a recall of 100%, resulting in an F1-Score of just 22.2%. Additionally, the word cloud analysis revealed that student complaints are prevalent, particularly regarding the system's slow performance, access errors to features, and challenges in filling out KRS and checking grades, highlighting the need for technical improvements.

With these enhancements, it is anticipated that the SIMAK system at Universitas Sriwijaya will become more responsive to student needs and provide a better user experience. Future research is recommended to explore other classification methods, such as support vector machines, for more accurate results, and to consider data from various platforms for a more comprehensive analysis.

Acknowledgement

Thank you to the Faculty of Computer Science for providing learning, writing scientific papers and research, especially the information systems department of the Faculty of Computer Science, Sriwijaya University

Reference

[1] S. Kemp, “Digital 2024: Indonesia,” Data Reportal. Accessed: Aug. 30, 2024. [Online]. Available: Document https://datareportal.com/reports/digital-2024-indonesia?rq=INDONESIA Table

[2] E. Sundari, “Transformasi Pembelajaran Di Era Digital: Mengintegrasikan Teknologi Dalam Pendidikan Modern,” *Sindoro Cendikia Pendidikan*, vol. 4, no. 4, pp. 25–35, 2024.

[3] H. Donan, E. S. N. S. Negara, T. Sutabri, and F. Firdaus, “Analysis of Behavioral Use of Academic Information Systems with the Implementation of UTAUT 2 Integration at the Muhammadi-Palembang Institute of Health Science and Technology,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 3, pp. 462–470, Nov. 2023, doi: 10.32736/sisfokom.v12i3.1978.

[4] S. Nurhanifah and D. R. Indah, “Evaluasi Kepuasan Pengguna pada Sistem Informasi Akademik (SIMAK) Universitas Sriwijaya dengan Menggunakan Metode PIECES Framework,” *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 11, no. 2, p. 339, Jul. 2023, doi: 10.26418/justin.v11i2.56574.

[5] M. Fajaria and K. Ditha Tania, “Evaluasi User Experience Dan Usability Sistem Informasi Akademik Menggunakan Metode User Experience Questionnaire Dan System Usability Scale,” *JOISIE (Journal Of Information System And Informatics Engineering)*, vol. 7, no. 2, pp. 204–213, Dec. 2023, doi: https://doi.org/10.35145/joisie.v7i2.3812.

[6] M. Fauzan Ahnaf, H. Muslimah Az-Zahra, W. Hayuhardhika, and N. Putra, “Pengukuran Kualitas Layanan Website Akademik Universitas Sriwijaya Kota Palembang Menggunakan Metode Webqual 4.0 dan Importance Performance Analysis (IPA),” *J-PTIIK*, vol. 3, no. 6, pp. 5817–5824, Jul. 2019.

[7] D. W. Ardras and A. Voutama, “Analisis Sentimen Anti Lgbt Di Indonesia Melalui Media Sosial Twitter,” *Jurnal Teknika*, vol. 15, no. 1, pp. 23–28, Mar. 2023, doi: 10.30736/jt.v15i1.926.

[8] C. Destitus, Wella, and Suryasari, “Support Vector Machine VS Information Gain: Analisis Sentimen Cyberbullying di Twitter Indonesia,” *Ultima InfoSys:Jurnal Ilmu Sistem Informasi*, vol. XI, no. 2, pp. 107–111, Dec. 2020, doi: https://doi.org/10.31937/si.v11i2.1740.

[9] L. Afuan, M. Khanza, and A. Z. Hasyati, “Enhancing Sentiment Analysis Of The 2024 Indonesian Presidential Inauguration On X Using Smote-Optimized Naive Bayes Classifier,” *Jurnal Teknik Informatika (JUTIF)*, vol. 6, no. 1, pp. 325–333, Feb. 2025, doi: 10.52436/1.jutif.2025.6.1.4290.

[10] I. Zulfahmi, J. Williem Iskandar, P. V Medan, S. Tuan, and D. Serdang, "Analisis Sentimen Aplikasi PLN Mobile Menggunakan Metode Decission Tree," *Jurnal Penelitian Rumpun Ilmu Teknik (JUPRIT)*, vol. 3, no. 1, pp. 11–21, 2024, doi: 10.55606/juprit.v3i1.3096.

[11] B. Ramadhani and R. R. Suryono, "Komparasi Algoritma Naïve Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *Jurnal Media Informatika Budidarma*, vol. 8, no. 2, p. 714, Apr. 2024, doi: 10.30865/mib.v8i2.7458.

[12] Ade Dwi Dayani, Yuhandri, and G. Widi Nurcahyo, "Analisis Sentimen Terhadap Opini Publik pada Sosial Media Twitter Menggunakan Metode Support Vector Machine," *Jurnal KomtekInfo*, pp. 1–10, Mar. 2024, doi: 10.35134/komtekinfo.v11i1.439.

[13] C. C. Dan and L. Saufa Yardha, "Analisis Sentimen Pengguna Aplikasi Mobile Jkn Menggunakan Algoritma Naïve Bayes," 2024. [Online]. Available: http://jurnal.goretanpena.com/index.php/JSSR

[14] T. V. Meiyanti, M. Hatta, and A. Sevtiana, "Analisis Sentimen Mahasiswa Dengan Dosen Menggunakan Metode Naïve Bayes Classifier Pada Kuesioner Dosen".

[15] I. Tri and J. J. Algoritma, "Analisis Sentimen Terhadap Sistem Informasi Akademik Mahasiswa Institut Teknologi Garut." [Online]. Available: https://jurnal.itg.ac.id/

[16] D. Lado Kaka, G. Kopong Pati, and K. Wulla Rato, "Analisis Sentimen Komentar SIAKAD Menggunakan Metode Naive Bayes Classifier." [Online]. Available: www.tripadvisor.com.

[17] L. R. Dharmawan, I. Arwani, and D. E. Ratnawati, "Analisis Sentimen pada Sosial Media Twitter Terhadap Layanan Sistem Informasi Akademik Mahasiswa Universitas Brawijaya dengan Metode K-Nearest Neighbor," 2020. [Online]. Available: http://j-ptiik.ub.ac.id

[18] I. Made, Y. Adhitya, D. E. Ratnawati, and I. Arwani, "Analisis Sentimen Feed Back Mahasiswa terhadap Dosen Prodi Teknologi Informasi Departemen Sistem Informasi Fakultas Ilmu Komputer Universitas Brawijaya," 2023. [Online]. Available: http://j-ptiik.ub.ac.id

[19] R. Fajar, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *Jurnal Inovtek Polbeng* , vol. 3, no. 1, pp. 50–59, 2020, doi: https://doi.org/10.35314/isi.v3i1.335.

[20] F. Nur Salsabilla and A. Witanti, "Analisis Sentimen Akhir Masa Jabatan Presiden Jokowi Pada Media Sosial X Menggunakan Naïve Bayes," *SKANIKA: Sistem Komputer dan Teknik Informatika*, vol. 8, no. 1, pp. 106–115, Jan. 2025, doi: https://doi.org/10.36080/skanika.v8i1.3331.

[21] N. C. Agustina, D. Herlina Citra, W. Purnama, C. Nisa, and A. Rozi Kurnia, "The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," *Machine Learning and Computer Science*, vol. 2, pp. 47–54, Apr. 2022, doi: https://doi.org/10.57152/malcom.v2i1.195.

[22] R. Asmara, M. Febrian Ardiansyah, and M. Anshori, "Analisa Sentiment Masyarakat terhadap Pemilu 2019 berdasarkan Opini di Twitter menggunakan Metode Naive Bayes Classifier," *Jurnal Inovtek Polbeng*, vol. 5, no. 2, pp. 193–204, 2020, doi: https://doi.org/10.35314/isi.v5i2.1095.

[23] R. A. Mrg and M. S. Hasibuan, "Best Student Classification using Ensemble Random Forest," *Sistemasi: Jurnal Sistem Informasi* , vol. 13, no. 3, pp. 1188–1204, 2024, doi: https://doi.org/10.32520/stmsi.v13i3.4101.

[24] M. Rizqi, A. Rinaldi Dikanda, and D. Rohman, "Applications Naive Bayes Algorithm to Enhance Sentiment Analysis of Coursera Application Reviews on Google Play Store," *Journal of Artificial Intelligence and Engineering*, vol. 4, no. 2, pp. 2808–4519, 2025, doi: https://doi.org/10.59934/jaiea.v4i2.758.

[25] L. Rofiqi and M. Akbar, "Analisis Sentimen Terkait RUU Perampasan Aset dengan Support Vector Machine," *JEKIN (Jurnal Teknik Informatika)*, vol. 4, no. 3, pp. 529–538, Aug. 2024, doi: 10.58794/jekin.v4i3.824.

[26] S. Abdullah, "Visualisasi Data Analisa Sentimen RUU Omnibus law Kesehatan Menggunakan KNN dengan Software RapidMiner," vol. 8, no. 3, 2023.

[27] M. Agym and H. Mayatopani, "Sentiment Analysis Public Perspection From Artemis 2 Mission Using Recurrent Neural Network Methods," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 4, pp. 385–393, 2024, doi: 10.52436/1.jutif.2024.5.4.2365.

[28] Y. Zaen Vebrian, "A Sentiment Analysis Of Free Meal Plans On Social Media Using Naïve Bayes Algorithms Analisis Sentimen Terhadap Rencana Makan Gratis Di Sosial Media X Menggunakan Alogaritma Naive Bayes," *Jurnal Inovtek Polbeng*, vol. 10, no. 1, p. 2025, 2025, doi: https://doi.org/10.35314/3m2fcz69.

[29] I. Habib Kusuma and N. Cahyono, "Analisis Sentimen Masyarakat Terhadap Penggunaan E-Commerce Menggunakan Algoritma K-Nearest Neighbor," *Jurnal Pengembangan IT (JPIT)*, vol. 8, no. 3, Sep. 2023, doi: https://doi.org/10.30591/jpit.v8i3.5734.

[30] A. Triyono, A. Faqih, and G. Dwilestari, "Implementation of the Naive Bayes Method in Sentiment Analysis of Spotify Application Reviews," *Journal of Artificial Intelligence and Engineering Applications* , vol. 4, no. 2, pp. 1091–1097, Feb. 2025, doi: https://doi.org/10.59934/jaiea.v4i2.824.

[31] Y. A. Singgalen, "Culture and heritage tourism sentiment classification through cross-industry standard process for data mining," *International Journal of Basic and Applied Science*, vol. 12, no. 3, pp. 110–120, Dec. 2023, doi: https://doi.org/10.35335/ijobas.v12i3.299.

[32] K. Kevin, M. Enjeli, and A. Wijaya, "Analisis Sentimen Pengunaaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes," *Jurnal Ilmiah Computer Science*, vol. 2, no. 2, pp. 89–98, Jan. 2024, doi: 10.58602/jics.v2i2.24.

[33] S. Ridwan Hidayat and Nurfaizah, "Sentimen Analisis Pengguna Produk Ponsel Menggunakan Algoritma Naïve Bayes," *Journal of Information System Management (JOISM) e-ISSN*, vol. 6, no. 1, pp. 2715–3088, 2024, doi: https://doi.org/10.24076/joism.2024v6i1.1625.

[34] M. I. Arif Chandra and R. Yusuf, "Visualisasi Kata Kunci Pemberitaan Pemilu 2024 Menggunakan Spacydan Wordcloud," *TEKNIMEDIA: Teknologi Informasi Dan Multimedia*, vol. 5, no. 1, pp. 41–46, 2024.

Assessing Academic Information System Performance Through Sentiment Analysis

ORIGINALITY REPORT

16%
SIMILARITY INDEX

PRIMARY SOURCES

1	sistemasi.ftik.unisi.ac.id Internet	42 words — 1%
2	Charlotte Nirmalani Gunawardena, Nick V. Flor, Damien M. Sánchez. "Knowledge Co-Construction in Online Learning - Applying Social Learning Analytic Methods and Artificial Intelligence", Routledge, 2025 Publications	29 words — 1%
3	jutif.if.unsoed.ac.id Internet	27 words — 1%
4	Sukhpreet Kaur, Sushil Kamboj, Manish Kumar, Arvind Dagur, Dhirendra Kumar Shukla. "Computational Methods in Science and Technology", CRC Press, 2024 Publications	26 words — 1%
5	Ruqiang Yan, Jing Lin. "Equipment Intelligent Operation and Maintenance", CRC Press, 2025 Publications	21 words — 1%
6	digibug.ugr.es Internet	21 words — 1%
7	Dewi Wardani, Chairul Achmad. "SSTI: Semantic Similarity to detect Novelty of Thesis Ideas",	18 words — 1%

Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications, 2022
Crossref

8	Citra Pricylia Ananda Mulya, Veri Arinal. "Analisis Sentimen Terhadap Sistem Informasi Akademik Mahasiswa pada Aplikasi Edlink dengan Metode K-Nearest Neighbor", Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi), 2024 Crossref	17 words — < 1%
9	Siti Nurmaini, Siti Zaiton, Dayang Norhayati. "An embedded interval type-2 neuro-fuzzy controller for mobile robot navigation", 2009 IEEE International Conference on Systems, Man and Cybernetics, 2009 Crossref	17 words — < 1%
10	Kazunori Iwata, Toyoshiro Nakashima, Yoshiyuki Anan, Naohiro Ishii. "Applying Machine Learning Classification to Determining Outliers in Effort for Embedded Software Development Projects", 2019 6th International Conference on Computational Science/Intelligence and Applied Informatics (CSII), 2019 Crossref	16 words — < 1%
11	S. Prasad Jones Christydass, Nurhayati Nurhayati, S. Kannadhasan. "Hybrid and Advanced Technologies", CRC Press, 2025 Publications	16 words — < 1%
12	Gaurav Aggarwal, Ashutosh Tripathi, Himani Goyal Sharma, Tripti Sharma, Rishabh Dev Shukla. "Integrated Technologies in Electrical, Electronics and Biotechnology Engineering", CRC Press, 2025 Publications	15 words — < 1%
13	Saepudin, Sutisna. "Sentiment Analysis of Kredivo App Users Using the K-Nearest Neighbor	15 words — < 1%

Algorithm", International Journal Software Engineering and Computer Science (IJSECS), 2024
Crossref

14	Tanrıkulu, Haluk. "Yapay Sinir Ağları Ile Bitki Hastalıklarının Teşhisi", Ankara Universitesi (Turkey), 2024 ProQuest	15 words — < 1%
15	idss.iocspublisher.org Internet	14 words — < 1%
16	Amit Kumar Tyagi, Shrikant Tiwari, Sayed Sayeed Ahmad. "Industry 4.0, Smart Manufacturing, and Industrial Engineering - Challenges and Opportunities", CRC Press, 2024 Publications	13 words — < 1%
17	Tasya Aisyah Amini, Kiki Setiawan. "Application of the Naive Bayes Algorithm in Twitter Sentiment Analysis of 2024 Vice Presidential Candidate Gibran Rakabuming Raka using Rapidminer", International Journal Software Engineering and Computer Science (IJSECS), 2024 Crossref	13 words — < 1%
18	Resti, Y. "Dependence in Classification of Aluminium Waste", Journal of Physics Conference Series, 2015. Crossref	12 words — < 1%
19	Khusnul Khotimah, Martanto Martanto, Arif Rinaldi Dikananda, Ahmad Rifa'i. "ANALISIS SENTIMEN ULASAN APLIKASI PINTU DI GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA NAIVE BAYES", Jurnal Informatika dan Teknik Elektro Terapan, 2025 Crossref	11 words — < 1%
20	ejournal.bsi.ac.id Internet	11 words — < 1%

21	ejournal.unsri.ac.id Internet	11 words — < 1%
22	ejurnal.seminar-id.com Internet	11 words — < 1%
23	Kartini, Aang Kisnu Darmawan, Riky Iskandar Syah, Masdukil Makruf. "Sentiment Analysis of Social Media for Indonesian m-Health PeduliLindungi Mobile-Apps(PLMA) with Lexicon-Based and Support Vector Machine Approach", 2023 IEEE 9th Information Technology International Seminar (ITIS), 2023 Crossref	10 words — < 1%
24	ejournal.insuriponorogo.ac.id Internet	10 words — < 1%
25	mdpi-res.com Internet	10 words — < 1%
26	Muhammad Saifurridho, Martanto Martanto, Umi Hayati. "Analisis Algoritma K-Nearest Neighbor terhadap Sentimen Pengguna Aplikasi Shopee", Jurnal Informatika Terpadu, 2024 Crossref	9 words — < 1%
27	doaj.org Internet	9 words — < 1%
28	ejeset.saintispub.com Internet	9 words — < 1%
29	hdl.handle.net Internet	9 words — < 1%
30	www.frontiersin.org Internet	9 words — < 1%

31	Agus Andreansyah, Julian Supardi. "Optimization of VGG-16 Accuracy for Fingerprint Pattern Imager Classification", Jurnal Sisfokom (Sistem Informasi dan Komputer), 2025 Crossref	8 words — < 1%
32	ejournal.pelitaindonesia.ac.id Internet	8 words — < 1%
33	ij.lafadzpublishing.com Internet	8 words — < 1%
34	injoit.ru Internet	8 words — < 1%
35	journal.irpi.or.id Internet	8 words — < 1%
36	jurnal.goretanpena.com Internet	8 words — < 1%
37	jurnalnasional.ump.ac.id Internet	8 words — < 1%
38	Du, Zhenjiao. "Machine Learning Empowered Discovery of Bioactive Peptides From Food Proteins and Beyond", Kansas State University, 2025 ProQuest	7 words — < 1%
39	Ali, W.. "Intelligent Naive Bayes-based approaches for Web proxy caching", Knowledge-Based Systems, 201207 Crossref	6 words — < 1%

EXCLUDE QUOTES OFF EXCLUDE SOURCES OFF

EXCLUDE BIBLIOGRAPHY ON EXCLUDE MATCHES OFF