J-CEKI: Jurnal Cendekia Ilmiah
Vol. 4, No. 6, Oktober 2025

Analysis of Public Perceptions of Tertiary Hospitals Through Google Maps Using Naïve Bayes and Wordcloud

Vira Aulia Kusuma Wardani[1], Arlette Suzy Puspa Pertiwi[2], Achmad Dheni Suwardhani[3]
[1,2,3]ARS University, Indonesia
E-mail: vira.fk19@hangtuah.ac.id[1], arlettesuzy@gmail.com[2], dheni88@gmail.com[3]

Article History:
Received: 30 Juli 2025
Revised: 05 September 2025
Accepted: 28 September 2025

Keywords: *Sentiment analysis, Naïve Bayes Classification, Wordcloud*

Abstract: *The study investigates the public perception of a tertiary hospital through Google Maps digital reviews using a Naïve Bayes approach and wordcloud. The data was collected by manual scraping technique from 2024-2025 and processed through preprocessing stages including cleansing, case folding, stopword removal, tokenizing, dan filter by legth. The data was then weighted by the TF-IDF technique before sentiment modeling by Naïve Bayes Classification. All of the processes use the Rapid Miner application. From 830 reviews collected, the accuracy of the Naïve Bayes model is 72,29% with precision 78,94%, recall 72,28%, and F1 score 74,76%. The results show that the dominant reviews are positive and describe the hospital health service as excellent. Therefore, the negative review shows that the hospital needs to improve its management in the pharmacy sector, administration, and patient queue. Neutral reviews tend to be descriptive without showing strong sentiments.*

INTRODUCTION

In recent years, there has been an increase in in the number of new hospitals. Data from 2010-2020 indicates a significant rise in the number of private hospitals in Indonesia, reflecting a growing trend within the healthcare sector (Rachmawati et al., 2024). And this event affects the competition among the hospitals. Healthcare institution needs to prioritize the delivery of High-quality services to ensure sustainability and meet the evolving needs of patients (Andriani et al., 2023; Darzi et al., 2023; Elizar et al., 2020).

One of the factors that influence a patient's decision to visit a hospital is to look at the hospital's public reviews. Indirectly, public reviews reflect the hospital's image, especially in the medical service provided (Soesilo et al., 2024). In this digital era, it is very easy to access this information (Andriani et al., 2022; Angel et al., 2024; Yassir et al., 2023). It is important for hospital to be able to understands and analyze these responses systematically (Utomo et al., 2023). This is where sentiment analysis comes in, by turning subjective opinions into objective data that can be used as the basis for strategic decision making (Wardani & Erfina, 2021).

Sentiment analysis helps analyse opinions along with the emotions contained to gain a deeper understanding of public perception of a topic (Makarim et al., 2024; Salsabila, 2022; Wardani & Erfina, 2021). Naive Bayes Classification (NBC) is one of the methods used in sentiment analysis. This NBC method is easy to do and fast in process, so it is very suitable for the sentiment analysis (Safira & Hasan, 2023).

LITERATURE REVIEW

Sentiment Analysis

Sentiment analysis or also called opinion mining is a field of study that analyzes a person's opinions, sentiments, evaluations, ratings, attitudes and emotions towards certain entities such as products, services, organizations, individuals, problems, events, topics, and others (Kusuma Wardani & Arum Sari, 2021). The basic task of sentiment analysis is to categorize the polarity of text in a document, sentence, or feature/aspect level whether the opinion expressed in the document, sentence or entity feature or aspect is positive, negative, or neutral (Lestari & Saepudin, 2021; Wankhade et al., 2022).

Sentiment analysis can come from various social media platforms such as Twitter, Facebook, product reviews, discussion forums, e-commerce, and others. The techniques and methods used in sentiment analysis are diverse and depend heavily on the type of data, the purpose of the analysis, and the available data sources (Huwaida et al., 2024; Putri et al., 2024; Trista & Asmoro, 2025). Some approaches that can be used include lexicon-based, machine learning, and deep learning approaches. Naïve Bayes classification belongs to the machine learning approach (Darwis et al., 2021; Ismail & Hakim, 2023; Makarim et al., 2024).

Sentiment analysis is widely applied in various sectors such as hospitality, airlines, healthcare and the stock market (Ruffer et al., 2020; Wankhade et al., 2022). However, sentiment analysis also has its own challenges, especially for reviews that are ambiguous, contain sarcastic sentences, or use local languages that are not included in the dictionary (Mikhael, 2022; Nugraha, 2017). Therefore, it is important to process review sentences into words that are easily understood and analyzed by the system in order to produce maximum accuracy (Trista & Asmoro, 2025).

Naïve Bayes Classification

Naïve Bayes classification is an algorithm that utilizes probability and statistical calculations proposed by an English scientist named Thomas Bayes. It predicts future probabilities based on previous experience. The Naïve Bayes algorithm is called "naïve" because it assumes that the occurrence of certain features is independent of the occurrence of other features. The Naïve Bayes classifier assumes that the value of a particular feature is dependent on the value of another feature, due to the class variable. The advantage of using this algorithm is that it requires very little training data to determine the parameter estimates required in the classification process. Since it is assumed to be an independent variable, it only requires the variance of a variable within a class to determine classification, not the entire covariance matrix (Saragih, 2021). The equation for calculating Naïve Bayes can generally be given as follows.

$$P(H|X)=\frac{(P(H \mid X)P(H)}{(P(X)}$$

Description:

X = Data with unknown class
H = Hypothesis that X is a specific class
P(H|X) = Probability of hypothesis H based on condition X
P(H) = Probability of hypothesis H
P(X) = Probability of X

Bayes' rule is as follows: If P(H1|X) < P(H2|X), then x is classified as H2. The statement P(H1|X) indicates the probability of hypothesis H1 based on the condition that X occurs, as well as H2 to classify X according to the largest probability among the probabilities of X to all classes (Saragih, 2021).

The accuracy of Naïve Bayes in classifying reviews varies and is influenced by various things, especially by the amount of training and test data used. In general, the greater the amount of training data used, the higher the accuracy compared to models trained using less training data (Gunawan et al., 2018; Rizki et al., 2025).

RESEARCH METHOD AND MATERIALS

Figure 1 shows the research flow. This research starts by doing web scraping manually by taking the review data and its rating. Any form of identity is not taken is this research. The data is then manually labeled by the researcher into positive, negative and neutral classes. After that, the data went through the preprocessing, model training, model evaluation, and word frequency (Wordcloud) stages. The entire analysis process was carried out using the Rapid Miner Instrument.

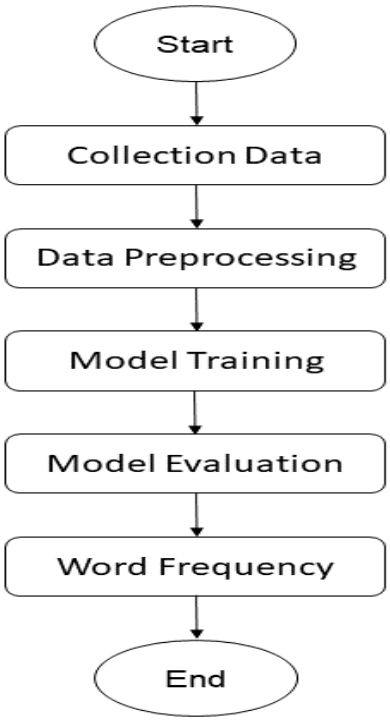

Figure 1. Research Flow

Data Preprocessing

The preprocessing stage is the first step in text processing which aims to clean and simplify the text to make it more ready for further analysis. The preprocessing stage involves several steps, namely cleansing, case folding, tokenization, stopword removal, and filter tokens by length (Y. Firmansyah et al., 2024; Sendi et al., 2023; Zamzami et al., 2024).

Word Weighting

Word weighting is a process to convert data from text form into numerical form, which can then be used as a basis for giving weight to assess how significant a word is in a document. By giving weight to each word that reflects how often and how significant the word appears in a data or document. This process involves applying a word weighting algorithm known as Term Frequency-Inverse Document Frequency (TF-IDF). Each word contained in the review will be

given a weight value generated through calculations using the TF-IDF algorithm (Y. Firmansyah et al., 2024; Zamzami et al., 2024).

Model Training

Model training is the process of creating a model that has intelligence based on the training data that has been given. Training data is obtained from 80% of the total review data that has been labeled and through preprocessing. The model used is the Naïve Bayes Classification model. Hopefully, the model is able to learn the training data and can classify the test data into positive, negative and neutral classes (Y. Firmansyah et al., 2024).

Model Evaluation

Model evaluation is an important stage in sentiment analysis to measure the extent to which the classification model that has been trained is able to make accurate predictions. The benchmarks used to measure performance are accuracy, precission, recall, and F1-Score. The higher the percentage value, the better the results of the model (Gunawan et al., 2018; Insan et al., 2023; Salsabila, 2022). If the data distribution is unbalanced, additional calculations can be performed using the weighted average formula. Weighted metrics adjust standard evaluation scores (such as precision, recall, and F1-score) by assigning weights proportional to class frequencies or importance, ensuring that minority classes are fairly represented in the overall performance measure instead of being overshadowed by dominant classes (Stout, 2025).

Word Frequency

Word cloud is a text mining method that graphically displays word frequencies that emphasize words that appear more frequently in the source text. The larger the size of the word in the visual, the more common the word is in the document (Limbong et al., 2022).

RESULTS AND DISCUSSION

The amount of data collected from the manual scraping process is 830 data, without any duplication data.

Data Preprocessing

The review data that has been collected is still unstructured, so it is necessary to process the data before analysis is caried out

1. Cleansing

At this stage, the review is cleaned of punctuation, numbers, special characters, double spaces and symbol characters ~!@#$%^&*()_+?<>,.?:{}[]. This is done using the “replace” ooperator in Rapid Miner by entering these characters in the “replace what” parameter and having the “replace by” field blank (Z. Firmansyah & Puspitasari, 2021; Putri et al., 2024).

2. Case folding

At this stage, capital letters in reviews are converted to lower case. In Rapid Miner, the Transform cases operator is set to “lower case” (Putri et al., 2024).

3. Stopwords removal

At this stage, connecting words in the review such as “di”, “ke”, “ini”, and “dari” are removed. This process uses the Filter Stopwords (Dictionary) operator of the Indonesian stopword dictionary file downloaded from the Kaggle website (https://www.kaggle.com/datasets/oswinrh/indonesian-stoplist) with the file name “stopwordbahasa.csv”. The file contains a list of common words in Indonesian that have no special

meaning in the analysis, so they can be removed from the data. By importing this file into the "file" parameter of the operator, all tokens corresponding to the stopword list will be automatically removed from the review data, allowing the analysis to focus on relevant and meaningful words (Putri et al., 2024).

4. Tokenizing

At this stage, the sentence is broken down into units of words or tokens. This stage is done using the Tokenize operator by changing the selected mode to "non letters" so that all characters other than letters are used as separators to break the text into words (tokens) (Z. Firmansyah & Puspitasari, 2021).

5. Filtering by length

In this process, words that are too short or abbreviated and words that are too long are removed. At this stage, the minimum value of characters (min charts) is set at 4 and the maximum characters (max charts) at 25. These settings ensure that only tokens or words that are between 4 and 25 characters long are retained in the data, while words that are too short or too long are removed. This step aims to improve data quality by focusing the analysis on more relevant and meaningful words (Putri et al., 2024).

Word Weighting

Word weighting is an important stage in sentiment analysis that aims to convert text data into numerical representations that can be processed by machine learning algorithms. The Process Documents from Data operator in RapidMiner for the feature extraction process uses the TF-IDF (Term Frequency-Inverse Document Frequency) method. At this stage, the "vector creation" option is set to TF-IDF so that each word in the review data will be converted into a numerical representation based on the weight of its occurrence in the document and the entire corpus. This setting allows the machine learning model to recognize and process text data more effectively, as words that are more important in the context of the document will get a higher weight (Y. Firmansyah et al., 2024; Zamzami et al., 2024).

Model Training

Reviews that have passed the preprocessing stage are divided into 80% and 20%. A total of 644 reviews were used as training data and 166 reviews as test data. The reviews are given to the Naïve Bayes model to form a model that has the intelligence to sort reviews into positive, negative and neutral classes. After forming the desired Naïve Bayes model, the model is tested using test data that has been previously prepared (Y. Firmansyah et al., 2024).

Model Evaluation

The outcome of the Naïve Bayes evaluation is a confusion matrix table containing the actual and predicted values of positive, negative, and neutral classes. The confusion matrix table is addressed by table 1.

Table 1. Confusion Matrix

	True Negative	True Neutral	True Positive
Prediction Negative	43	6	12
Prediction Neutral	1	0	18
Prediction Positive	6	3	77

From the table above, we get the actual data distribution:
Positive : 107 data (12+18+77)
Negative: 50 data (43+1+6)
Neutral: 9 data (6+0+3)
Total: 166 data

Details of True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) values for each sentiment class:

1. Positive Class
   a. True Positive : 77 (predicted as positive and true positive)
   b. False Positive : 9 (predicted positive but should have been negative/neutral)
   c. False Negative : 30 (should be positive but predicted to be negative/neutral)
   d. True Negative : 50 (not predicted to be positive and not actually positive)
2. Negative Class
   a. True Positive : 43 (predicted as negative and true negative)
   b. False Positive : 18 (predicted negative but should have been positive/neutral)
   c. False Negative: 7 (should be negative but predicted to be positive/neutral)
   d. True Negative : 98 (not predicted to be negative and not actually negative)
3. Neutral Class
   a. True Positive : 0 (predicted as neutral and true neutral)
   b. False Positive : 19 (predicted neutral but should have been negative/positive)
   c. False Negative: 9 (should be neutral but predicted to be negative/positive)
   d. True Negative : 138 (not predicted to be neutral and not actually neutral)

After determining the confusion matrix of the model, the accuracy value of the Naive Bayes classification model is calculated.

1. Accurary

$$Accuracy = \frac{TPtotal}{Total\ Data} = \frac{43+0+7}{166} = \frac{120}{166} = 72{,}29\%$$

Description:
TP total: Total value of True Positive for all classes
Total data: Total test data

2. Precission

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

   a. Positive: $\frac{77}{77+9} = \frac{77}{86} = 89{,}53\%$
   b. Negative: $\frac{43}{43+18} = \frac{43}{61} = 70{,}49\%$
   c. Neutral: $\frac{0}{0+19} = \frac{0}{19} = 0{,}00\%$

3. Recall

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

   a. Positive: $\frac{77}{77+30} = \frac{77}{107} = 71{,}96\%$
   b. Negative: $\frac{43}{43+7} = \frac{43}{50} = 86{,}00\%$

c. Neutral: $\frac{0}{0+9}=\frac{0}{9}=0{,}00\%$

4. F1-Score

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

a. Positive: $\frac{2\times 89{,}53\times 71{,}96}{89{,}53+71{,}96}=79{,}82\%$

b. Negative: $\frac{2\times 70{,}49\times 86{,}00}{70{,}49+86{,}00}=77{,}42\%$

c. Neutral: $\frac{2\times 0{,}00\times 0{,}00}{0{,}00+0{,}00}=0{,}00\%$

5. Weighted Average

$$\sum_{i=1}^{n}\frac{(Metric_i\times Support_i)}{Total\ Support}$$

a. Weighted Precission : $\frac{(70{,}49\times 50)+(0{,}00\times 9)+(89{,}53\times 107)}{166}=\frac{3254{,}5+0+9579{,}71}{166}=78{,}94\%$

b. Weighted Recall : $\frac{(86{,}00\times 50)+(0{,}00\times 9)+(71{,}96\times 107)}{166}=\frac{4300+0+7699{,}72}{166}=72{,}28\%$

c. Weighted F1-Score : $\frac{(77{,}42\times 50)+(0{,}00\times 9)+(79{,}82\times 107)}{166}=\frac{3871+0+8640{,}74}{166}=74{,}76\%$

Word Frequency

Word frequency was performed using the Wordcloud feature in Rapid Miner. Figures 4, 5, and 6 show the visualization results of words in positive, negative, and neutral reviews.

Figure 2. Word cloud of positive reviews

Based on the analysis results, the most dominant word is “service” (302 times), followed by “friendly” (200 times), “doctor” (117 times), ‘good’ (107 times), and “nurse” (104 times). Regarding the reviews frequently discussed in positive reviews, aspects such as service, friendly attitude, and the quality of doctors and nurses are the main topics most appreciated by patients in positive reviews. Additionally, words like “thank you,” “good,” and “service” also frequently appear, indicating patient satisfaction and gratitude toward the hospital's services.

Figure 3. Word cloud of negative reviews

Based on the analysis results, the most frequently occurring word is “patient” (183 times), followed by “service” (101 times), “doctor” (99 times), “medicine” (87 times), “illness” (82 times), “BPJS” (66 times), “nurse” (63 times), ‘treatment’ (63 times), and “queue” (60 times). In negative reviews, the dominance of the words that appear indicates that the main complaints of patients in negative reviews are related to service, patient handling, medication availability, the BPJS system, as well as the treatment process and queues. This is in line with the findings of other studies which mention that issues with queue management, the attitude of medical staff, and inadequate facilities and administrative services are the aspects most frequently complained about in negative hospital reviews.

Figure 4. Word cloud of neutral reviews

Based on the analysis results, the most frequently occurring word is “service” (24 times), followed by “patient” (13 times), “good” (13 times), “room” (13 times), “friendly” (10 times), ‘nurse’ (8 times), and “hospital” (8 times). In neutral reviews, the words that appear indicate that the discussion still focuses heavily on aspects of service, patients, and room facilities, but without strong positive or negative sentiment. Additionally, words like “friendly” and “nurse” also appear, but the context tends to be descriptive or informative, not containing praise or complaints.

Discussion

This sentiment analysis yielded an accuracy of 72.29%. This is considered sufficient to meet the “good” criterion in the Naive Bayes approach, which is 70% (Ahmad et al., 2021). Several other sources set the minimum accuracy value for “good” at 75-85%. Accuracy values are influenced by various factors, especially the amount of training data and test data. The higher the ratio of training data to test data, the higher the accuracy value. In Gunawan's (2018) study, a ratio of 80:20 resulted in an accuracy of 52.66%, while using a ratio of 90:10 resulted in an accuracy of 59.33%. A larger number of classes can reduce accuracy compared to fewer classes. Testing with 5 classes has lower accuracy than 3 classes (Gunawan et al., 2018). Other analysis methods have higher accuracy values than Naïve Bayes; however, when considering the data format, the review classification process achieves higher accuracy with the Naïve Bayes method (Suhadi, 2025).

Regarding the precision, recall, and F1-Score values for the neutral class, which are only 0%, this is due to the insufficient amount of training and test data used. The selection of training data samples did not specify a data quota per class, resulting in an uneven distribution of data. Additionally, reviews in the neutral class have a broad vocabulary, often leading to misclassification into other review classes (Gunawan et al., 2018).

CONCLUSION

For Conclusions, in general, the analysis of public perception of tertiary hospitals through Google Maps digital reviews using the Naive Bayes and Wordcloud approaches yielded predominantly positive results with an accuracy of 72.29%. Positive reviews primarily focus on the quality of service and the attitude of medical staff. Negative reviews highlight complaints related to the service process, medication availability, as well as issues with queues and administration. In neutral reviews, the frequently appearing words are more descriptive in nature and do not indicate strong sentiment.

REFERENCES

Andriani, R., Andriani, P., & Andriani, D. (2023). Company performance improvement: implementation of service culture through human capital in hospitality industry. *INOVASI: Jurnal Ekonomi, Keuangan, Dan Manajemen*, *19*(1), 10–16.

Andriani, R., Veranita, M., Noor, C. M., Ismail, K., Fauzzia, W., Hidayat, F. A., & Parino, N. H. S. (2022). Strategi Peningkatan Pemasaran Melalui Digital Marketing Pada Umkm Binangkit Kabupaten Bandung. *JPKMI (Jurnal Pengabdian Kepada Masyarakat Indonesia)*, *2*(3), 356–359.

Ahmad, A., Sakidin, H., Sari, M. Y. A., Amin, A. R. M., Sufahani, S. F., & Rasib, A. W. (2021). Naïve Bayes Classification of High-Resolution Aerial Imagery. (IJACSA) International Journal of Advanced Computer Science and Applications, 12(11), 168–177.

Angel, A. C. T., Pranatawijaya, V. H., & Widiatry, W. (2024). Analisis Sentimen dan Emosi dari Ulasan Google Maps Untuk Layanan Rumah Sakit di Palangka Raya Menggunakan Machine Learning. KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi, 4(1), 35–49. https://doi.org/10.24002/konstelasi.v4i1.8924

Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes untuk Analisis Sentimen Review Data Twitter BMKG Nasional. Jurnal TEKNO KOMPAK, 15(1), 131–135. https://doi.org/https://doi.org/10.33365/jtk.v15i1.744

Darzi, M. A., Islam, S. B., Khursheed, S. O., & Bhat, S. A. (2023). Service quality in the healthcare sector: a systematic review and meta-analysis. LBS Journal of Management & Research, 21(1), 13–29. https://doi.org/10.1108/LBSJMR-06-2022-0025

Elizar, C., Indrawati, R., & Syah, T. Y. R. (2020). Service Quality, Customer Satisfaction, Customer Trust, and Customer Loyalty in Service of Paediatric Polyclinic Over Private H Hospital of East Jakarta, Indonesia. Journal of Multidisciplinary Academic, 4(2).

Firmansyah, Y., Kurniawan, R., & Wijaya, Y. A. (2024). Analisis Data Sentimen Pemain Game Role-Playing Game (RPG) Honkai Star Rail dengan Algoritma Naive Bayes. Jurnal Informatika Dan Rekayasa Perangkat Lunak, 6(1), 127–135. https://sistemasi.ftik.unisi.ac.id/index.php/stmsi/article/download/4343/799

Firmansyah, Z., & Puspitasari, N. F. (2021). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 Berdasarkan Opini Pada Twitter Menggunakan Algoritma Naive Bayes. Jurnal Teknik Informatika, 14(1), 171–178. https://doi.org/https://doi.org/10.15408/jti.v14i2.24024

Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. JEPIN (Jurnal Edukasi Dan Penelitian Informatika), 4(2), 113–118.

Huwaida, S. F., Kusumawati, R., & Isnaini, B. (2024). Analisis sentimen komentar youtube terhadap pemindahan ibu kota negara menggunakan metode Naïve Bayes. Jambura Journal of Informatics, 6(1), 26–39. https://doi.org/https://doi.org/10.37905/jji.v6i1.24718

Insan, M. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes. JATI (Jurnal Mahasiswa Teknik Informatika), 7(1), 478–483. https://doi.org/10.36040/jati.v7i1.6373

Ismail, A. R., & Hakim, R. B. F. (2023). Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Mengetahui Trend Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter. Emerging Statistics and Data Science Journal, 1(1), 37–46. https://doi.org/https://doi.org/10.20885/esds.vol1.iss.1.art5

Kusuma Wardani, S., & Arum Sari, Y. (2021). Analisis Sentimen menggunakan Metode Naïve

Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan Seleksi Fitur N-gram dan Document Frequency Thresholding (Vol. 5, Issue 12). http://j-ptiik.ub.ac.id

Lestari, S., & Saepudin, S. (2021). Analisis Sentimen Vaksin Sinovac pada Twitter Menggunakan Algoritma Naïve Bayes. SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika) , 1(7), 163–170. https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/23/20

Limbong, J. J. A., Sembiring, I., & Hartomo, K. D. (2022). Analisis Klasifikasi Sentimen Ulasan Pada E-Commerce Shopee Berbasis Word Cloud Dengan Metode Naive Bayes Dan K-Nearest Neighbor. Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIIK), 9(2), 347–356.

Makarim, Z., Nawangsih, I., & Sanudin, S. (2024). Naive Bayes Algorithm for Sentiment Analysis on Spider-Man Movie: No Way Home. Journal of Computer Networks, Architecture and High Performance Computing, 6(4), 1886–1897. https://doi.org/10.47709/cnahpc.v6i4.4845

Mikhael, L. (2022). Sosialisasi Pemahaman Larangan Undang-Undang Informasi Dan Transaksi Elektronik Serta Etika Penggunaan Media Sosial Pada Remaja. AIWADTHU: Jurnal Pengabdian Hukum, 2(2), 50. https://doi.org/10.47268/aiwadthu.v2i2.940

Nugraha, A. L. (2017). Pendeteksian Ambiguitas Makna Kata Untuk Meningkatkan Performa Analisa Sentimen Dengan Menggunakan Algoritma Similaritas Path Length. Universitas Muhammadyah Surakarta.

Putri, K. S., Setiawan, I. R., & Pambudi, A. (2024). Analisis Sentimen Terhadap Brand Skincare Lokal Menggunakan Naïve Bayes Classifier. Technologia, 14(3), 227–232. https://doi.org/http://dx.doi.org/10.31602/tji.v14i3.11259

Rachmawati, E., Yuyun Umniyatun, Deni Wahyudi Kurniawan, Mochamad Iqbal Nurmansyah, Mukhaer Pakkanna, Husnan Nurjuman, Slamet Budiarto, & Virgo Sulianto Gohardi. (2024). Analysis Of The Market Structure Of Hospital Industry In Indonesia. Jurnal Administrasi Kesehatan Indonesia, 12(1), 37–48. https://doi.org/10.20473/jaki.v12i1.2024.37-48

Rizki, A. F., Prihartono, W., & Rohman, F. (2025). Analisis Sentimen Ulasan Google Maps Rumah Sakit Khalishah di Cirebon dengan Algoritma Naive Bayes. Jurnal Informatika Dan Teknik Elektro Terapan, 13(2), 418–427. https://doi.org/10.23960/jitet.v13i2.6309

Ruffer, N., Knitza, J., & Krusche, M. (2020). #Covid4Rheum: an analytical twitter study in the time of the COVID-19 pandemic. Rheumatology International, 40(12), 2031–2037. https://doi.org/10.1007/s00296-020-04710-5

Safira, A., & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes. Jurnal Sistem Informasi, 5(1), 59–70. https://doi.org/https://doi.org/10.31849/zn.v5i1.12856

Salsabila, N. A. (2022). Analisis Sentimen Pada Media Sosial Twitter Terhadap Tokoh Gus Dur Menggunakan Metode Naïve Bayes dan Support Vector Machine (SVM). In Universitas Islam Negeri Syarif Hidayatullah. Universitas Islam Negeri Syarif Hidayatullah. https://repository.uinjkt.ac.id/dspace/handle/123456789/66253

Saragih, H. M. (2021). Analisis Sentimen Pengguna Twitter Terhadap Layanan Pajak Kendaraan Bermotor Menggunakan Algoritme Naive Bayes Classifier. In Universitas Lampung. Universitas Lampung.

Sendi, A. R., Alam, S., & Kurniawan, I. (2023). Analisis Sentimen Pengguna Aplikasi Jmo (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes. 2(3),

109–117. https://doi.org/https://doi.org/10.55123/storage.v2i3.2334

Stout, D. W. (2025, February 21). Weighted Metrics for Multi-Class Models Explained. Magai. https://magai.co/weighted-metrics-for-multi-class-models-explained/

Suhadi, E. (2025). Analisis Sentimen Aplikasi Bisa Ekspor Pada Ulasan Pengguna Di Google Play Dengan Naïve Bayes. JIKA (Jurnal Informatika) Universitas Muhammadiyah Tangerang, 9(1), 93–101.

Soesilo, D., Rohendi, A., Hidayat, D., & Syaodih, E. (2024). Efektivitas Pemasaran Digital, Citra Merek, Variabel Intervening Kepercayaan Pasien Pada Rsgmp Nala Husada Surabaya. *Journal of Innovation Research and Knowledge*, *4*(6), 3807–3816.

Trista, R. T., & Asmoro, E. T. (2025). Analisis Sentimen dalam Data Big Data (Studi Kasus pada Media Sosial). Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research , 2(1), 700–706. https://doi.org/https://doi.org/10.32672/mister.v2i1.2518

Utomo, M. A., Purwadhi, P., & Hidayat, D. (2023). Pengaruh Manajemen Sarana Kesehatan Terhadap Kepuasan Pasien Di Klinik Pratama Ptpn Vii Kantor Direksi Bandar Lampung. *Referensi : Jurnal Ilmu Manajemen Dan Akuntansi*, *11*(2), 107–116. https://doi.org/10.33366/ref.v11i2.4641

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Wardani, N. R., & Erfina, A. (2021). Analisis Sentimen Masyarakat Terhadap Layanan Konsultasi Dokter Menggunakan Algoritma Naive Bayes. SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika), 11–18.

Yassir, A., Purwadhi, P., & Andriani, R. (2023). Hubungan mutu pelayanan terhadap minat kunjungan ulang pasien di klinik citra medika kota Semarang. *Jurnal Riset Pendidikan Ekonomi*, *8*(1), 1–12. https://doi.org/10.21067/jrpe.v8i1.8239

Zamzami, F., Hidayat, R., & Fathonah, R. (2024). Penerapan Algoritma Naive Bayes Classifier Untuk Analisis Sentimen Komentar Twitter Proyek Pembagunan IKN. Faktor Exacta, 17(1). https://doi.org/10.30998/faktorexacta.v17i1.22265