J-INTECH (Journal of Information and Technology)
Accredited Sinta 4 Ministry of Higher Education, Science and Technology
Republic of Indonesia SK No. 10/C/C3/DT.05.00/2025
E-ISSN: 2580-720X || P-ISSN: 2303-1425

Application of Naive Bayes Algorithm for Analysis of User Reviews on Mobile Legends Game: Bang Bang

Al-Muchlis Syachrul Ramadani Roba[1*], Siti Lailiyah[2], Amelia Yusnita[3]

[1,2]STMIK Widya Cipta Dharma, Informatics Engineering, Jl. M. Yamin No.25, Gn. Kelua, Kec. Samarinda Ulu, Kota Samarinda, Kalimantan Timur 75123, Indonesia
[3]STMIK Widya Cipta Dharma, Information Systems, Jl. M. Yamin No.25, Gn. Kelua, Kec. Samarinda Ulu, Kota Samarinda, Kalimantan Timur 75123, Indonesia

Keywords

Mobile Legends; Naïve Bayes; Sentiment Analysis.

****Corresponding Author:***
2143029@gmail.ac.id

Abstract

Mobile Legends: Bang Bang is a highly popular MOBA game, especially among students, which generates a large volume of user reviews on the Google Play Store. These reviews provide a valuable data source for understanding user sentiment. This study conducts sentiment analysis on user reviews using three variants of the Naïve Bayes algorithm: BernoulliNB, GaussianNB, and MultinomialNB. From an initial 5,000 reviews collected via web scraping using Python, 4,428 reviews were used after neutral reviews were removed to focus solely on positive and negative sentiments.The preprocessing steps included case folding, word normalization, tokenization, stopword removal, and stemming. Sentiment labeling was carried out using a lexicon-based approach, comparing the frequency of positive and negative words in each review. The dataset was split in an 80:20 ratio for training and testing.The results show that MultinomialNB achieved the highest accuracy at 75%, followed by BernoulliNB with 74%, and GaussianNB with 50%. MultinomialNB demonstrated superior performance in detecting positive sentiments, while BernoulliNB offered more balanced results. GaussianNB performed poorly due to its assumption of normally distributed continuous data, which is unsuitable for text classification. This study concludes that Multinomial Naïve Bayes is the most effective model for sentiment analysis of user reviews when working with word frequency-based representations.

1. Introduction

Online games come in various genres such as Puzzle, RPG, and MOBA. MOBA games, such as *Mobile Legends*, are popular because they involve two teams competing against each other to destroy the opponent's base. *Mobile Legends* is played by five players per team, each controlling a hero, assisted by minions, and fighting in three lanes (top, middle, bottom). The game also provides items such as heroes and skins. Despite its popularity among students, the game has raised concerns due to its negative impact on physical, mental and social health. However, online games also have positive impacts such as reducing stress and improving English language skills [1].

Through the comment page of the *Mobile Legends: Bang Bang* game found on the Play Store, we can observe user reviews to assess customer satisfaction. To analyze these reviews, researchers use data mining techniques, particularly sentiment analysis. The field of sentiment analysis is growing rapidly as more users leave online reviews. Sentiment analysis is commonly used to evaluate public opinion by analyzing text data and classifying it into categories such as positive or negative sentiment [11].

This research uses the Naïve Bayes algorithm, a classification method based on probability [4]. Prediction calculations in this study will utilize the Naïve Bayes algorithm. Previous studies have found that Naïve Bayes offers a relatively good level of accuracy [13]. However, the effectiveness of Naïve Bayes in complex sentiment contexts remains questionable. For example, in a prior study comparing Naïve Bayes and SVM for sentiment analysis on Twitter data related to the corruption of social aid rice during the pandemic, Naïve Bayes achieved only 60.61% accuracy, while SVM reached 66.67%. This suggests that Naïve Bayes may struggle to capture nuanced or context-rich sentiments.

Therefore, this research aims to analyze user sentiment towards the game *Mobile Legends: Bang Bang* based on Google Play Store reviews, while also evaluating the performance of the Naïve Bayes algorithm in classifying reviews into positive or negative categories. In doing so, this study highlights the need to reassess the reliability of Naïve Bayes and potentially explore alternative or complementary methods for more accurate sentiment classification.

2. Research Method

Methods To conduct this research, it is necessary to know how data collection procedures and data processing stages are carried out in data preprocessing, and classification modeling using the naive bayes classifier algorithm. The stages for the research method can be seen in Figure 1 below :

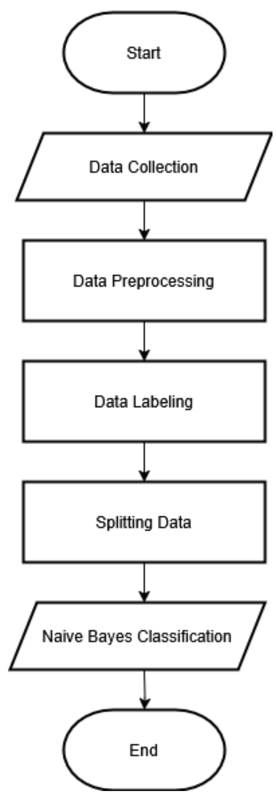


Figure 1. Research process flowchart

Data collection or collection of data to be taken from Google Play Stored using web screping techniques. Web scraping is a method of automatically retrieving information from websites and data is extracted[4]. Data preprocessing is conducted to prepare the data for modeling. Preprocessing is one of the techniques in data mining to convert raw data into a more understandable format. Preprocessing aims to remove noise and non-uniform word forms to reduce vocabulary volume[5]. At this stage there are several steps that must be taken such as data claning, case folding, word normalization, tokenizing, stopword removal, stemming.

Labeling in sentiment analysis is the process of classifying data based on sentiment categories such as positive and negative. This process helps machine learning models understand and group data based on predetermined categories, and in this study, I used Lexicon-based labeling. Lexicon-based labeling is often preferred over manual annotation or supervised methods because it offers a fast, cost-effective, and scalable way to label large volumes of text without requiring human effort or pre-labeled training data. Unlike manual annotation, which is time-consuming and subject to human bias or inconsistency, lexicon-based approaches rely on predefined sentiment dictionaries to classify text automatically. This makes them particularly useful for initial sentiment labeling or when resources are limited. While less accurate than supervised learning and unable to capture context, sarcasm, or domain-specific nuances, lexicon-based methods are a practical starting point for many sentiment analysis tasks. [6].

In this study, reviews with neutral scores (equal number of positive and negative words) were removed from the dataset at the labeling stage. This was done because neutral reviews are difficult to distinguish from positive reviews in the lexicon-based approach used, which risks compromising the accuracy of the model if they are forced into either class. In addition, adding a neutral class may complicate the model without providing significant analytical value. Therefore, only reviews that have a clear tendency towards positive or negative sentiment are used in training and testing the model.

Splitting data is the process of dividing a dataset into two or more separate parts, namely into training data and testing data [7]. The purpose of this splitting is to allow the model to learn from the training data and then measure its accuracy using the testing data, so that it can be assessed how well the model predicts new data. One of the commonly used methods for this data split is the Hold-Out method, which splits the data into two parts, where one part is used as training data and the other part is used as testing data.

This stage focuses on the software from a logical and functional point of view and ensures that all parts have been tested. The confusion matrix provides discrimination results with more incorrect and correct prediction data than the facts obtained. Accuracy, precision and F1 score are used. F1 score is a weighted average score of gain and precision, and accuracy measures how often the classification model correctly determines the correct class[8].

3. Result and Discussions

The data collection stage in this research was carried out by applying web scraping techniques, which is a method of automatically retrieving data from websites. This process uses the Python programming language run through the Google Collaboratory platform, which allows for the execution of code online with high efficiency. In this research, 5,000 user review entries from the Shopee app on the Google Play Store were collected.

The web scraping technique works by browsing the HTML document structure of a website to obtain relevant information. The information is then identified and retrieved for use as needed. In this context, the collected review data is analyzed and further processed for research purposes, particularly in sentiment analysis. This web scraping process refers to the approach described by[9], which states that web scraping is an effective method of retrieving user review data from platforms such as the Google Play Store using Python.

```
[6] data.head(5)
```

	Review ID	Username	Rating	Review Text	Date
0	2f7f3d91-c999-4460-a312-ddbc174fa126	Pengguna Google	1	Entah kenapa belakangan ini jaringan sering bg...	2025-05-01 02:03:10
1	47cc3eeb-96c8-44a8-a00b-4b7b21bbdbd5	Pengguna Google	3	bagus, tapi sering tiba-tiba bug sama ga tersa...	2025-05-01 07:26:46
2	1ed4acb0-1711-464d-b5bf-dab0e5a38d28	Pengguna Google	4	game yg bisa bikin user nya darah tinggi, kare...	2025-04-27 13:24:30
3	097ac278-d594-4a84-9113-fd4be4815aa0	Pengguna Google	5	apalahhh moonton ini,WS 7×,eh Kalah 10×,apalah...	2025-04-27 04:03:29
4	a4db583f-64f5-42e4-8ade-b3b3ef2bfd20	Pengguna Google	4	monton ini gimana ya? padahal udah di update m...	2025-04-30 05:21:16

Figure 2. Scraper data result

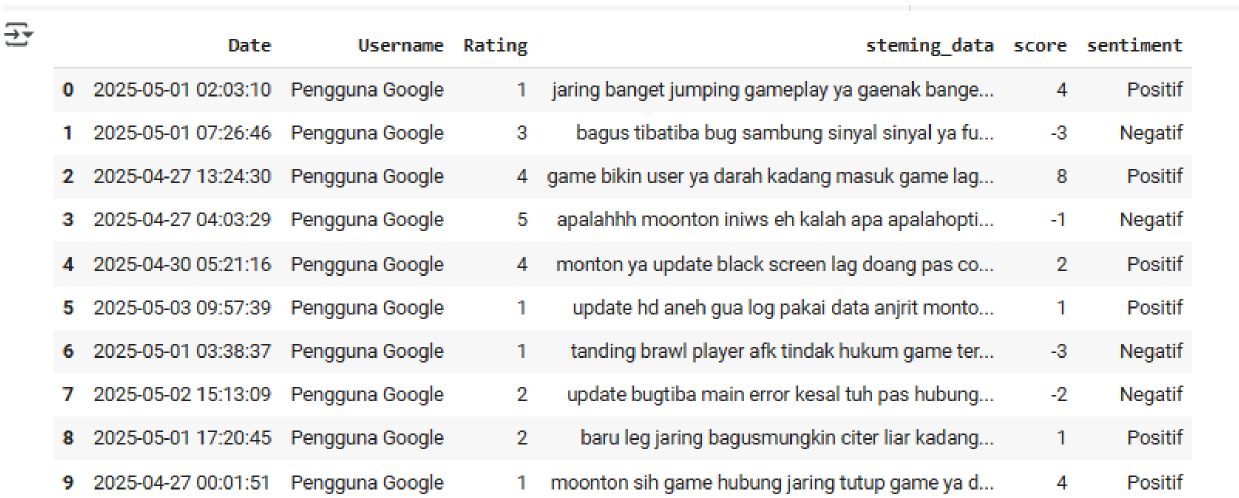
The collected review data is text data in .csv format and is still unstructured. Unstructured data can hinder the smoothness of the analysis process, so a text preprocessing stage is needed to turn the data into more structured and ready for processing. In the text preprocessing process, there are four main stages carried out, namely case folding, tokenizing, filtering, and stemming. These stages aim to clean and simplify text data so that it can be used in further analysis processes such as classification or sentiment analysis.

Case folding is a step in text preprocessing that aims to convert all letters in the review text into lowercase letters. This step is important to ensure consistency in text analysis, especially in avoiding discrepancies between words that are identical but written with different capital letters, such as "Application" and "app". By performing case folding, we can simplify the text and reduce unnecessary variations, so that sentiment analysis can be performed more accurately. Here is an example of the case folding results on the review dataset[10].

Word normalizer is a technique in text processing to simplify words that have the same meaning into one base form. For example, the words “fix” and ‘repair’ are normalized to “good.” According to Fajri , word normalization is important to improve the accuracy of text analysis, especially in processing user review data[11]. The tokenization process is an important step in text data processing. In this process, the text is broken down into small parts called tokens, which are the smallest meaningful units, such as words, numbers, or punctuation marks. The results of the tokenization process can be seen in the following figure[7]

Stages that function to stop words that are considered common words, pronouns, and conjunctions. For example, "Andi often conducts routine transactions online. According to Andi, online shopping is more practical & cheap“ will be changed to ”Andi often does routine transactions online or online, also Andi online shopping is more practical and cheap[12]. Stemming is to find the root word or the beginning of words that have affixes. This process is done by removing affixes, such as prefixes, suffixes, and infixes, so as to produce basic words that have the same meaning. The results of the stemming process in this research use the Sastrawi library on Indonesian text.[13]

Data labeling in this study uses the lexicon-based method, which is a dictionary-based approach of words that have been classified as positive or negative. Each review in the `steming_data` column is analyzed by counting the number of positive and negative words based on two lexicons: `positive.tsv` and `negative.tsv`. The difference of the number of positive and negative words yields the sentiment score. If the score > 0 then the review is labeled “Positive”, if the score < 0 then it is labeled “Negative”. When the score = 0, the label is randomly determined to be either positive or negative. This method was chosen because it is simple, does not require training data, and is suitable for initial text-based sentiment analysis.



	Date	Username	Rating	steming_data	score	sentiment
0	2025-05-01 02:03:10	Pengguna Google	1	jaring banget jumping gameplay ya gaenak bange...	4	Positif
1	2025-05-01 07:26:46	Pengguna Google	3	bagus tibatiba bug sambung sinyal sinyal ya fu...	-3	Negatif
2	2025-04-27 13:24:30	Pengguna Google	4	game bikin user ya darah kadang masuk game lag...	8	Positif
3	2025-04-27 04:03:29	Pengguna Google	5	apalahhh moonton iniws eh kalah apa apalahopti...	-1	Negatif
4	2025-04-30 05:21:16	Pengguna Google	4	monton ya update black screen lag doang pas co...	2	Positif
5	2025-05-03 09:57:39	Pengguna Google	1	update hd aneh gua log pakai data anjrit monto...	1	Positif
6	2025-05-01 03:38:37	Pengguna Google	1	tanding brawl player afk tindak hukum game ter...	-3	Negatif
7	2025-05-02 15:13:09	Pengguna Google	2	update bugtiba main error kesal tuh pas hubung...	-2	Negatif
8	2025-05-01 17:20:45	Pengguna Google	2	baru leg jaring bagusmungkin citer liar kadang...	1	Positif
9	2025-04-27 00:01:51	Pengguna Google	1	moonton sih game hubung jaring tutup game ya d...	4	Positif

Figure 3. Data Labeling Result

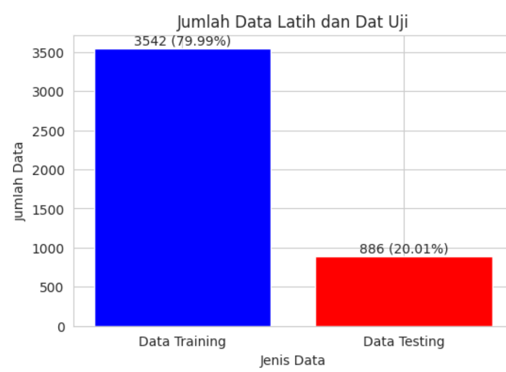


Figure 4. Spliting data result

This visualization shows the division of 4,428 review data into training and testing data after the splitting process. The bar graph displays the amount of each data, along with its percentage. The training data is shown in blue, while the testing data is red. The purpose of this visualization is to ensure that the proportion of data split, e.g. 80% for training and 20% for testing, is appropriate before the model is built.The initial amount of data was 5,000 reviews, but was reduced to 4,428 reviews after neutral reviews were removed. This was done to focus the analysis only on positive and negative reviews.

Table 1. Naive Bayes Classification

Model Classification	Percision		Recal		F1-score		Acurasi
	negatif	positif	negatif	positif	negatif	positif	
GaussianNB	46%	62%	77%	29%	58%	53%	50%
MultinomialNB	86%	71%	52%	93%	65%	81%	75%
BernouliNB	74%	74%	63%	82%	68%	78%	74%

The table shows the performance comparison of three Naive Bayes classification models, namely GaussianNB, MultinomialNB, and BernoulliNB, based on the evaluation metrics: precision, recall, f1-score, and accuracy. In general, GaussianNB has the lowest performance compared to the other two models. Its precision is only 47% for the negative class and 62% for the positive, with an unbalanced recall: high in negative (77%) but very low in positive (29%). This shows that GaussianNB tends to recognize negative data more often, but fails to capture positive data well. Its accuracy was only 51%, almost equivalent to random guessing.

In contrast, MultinomialNB showed the best performance. It had high precision (86% negative, 71% positive), as well as very strong recall in the positive class (93%), although the negative recall was lower (52%). F1-score was also high in both classes, and overall accuracy reached 75%. This shows that MultinomialNB is very good at detecting positive data and is quite robust in general.

Meanwhile, BernoulliNB gave balanced results between the two classes. Precision (74% for negative and 74% for positive) recall(63% for negative and 82% for positive), with an evenly high f1-score (68% for negative and 78% for positive). The total accuracy was 74%, only slightly below MultinomialNB, but with a better balance between the negative and positive classes. Thus, if the main objective is to recognize maximum positive data, MultinomialNB is the best choice. However, if a balance in detecting both classes equally is required, BernoulliNB can be an excellent alternative.

MultinomialNB performs best because this model is specifically designed for discrete data such as word frequency in text. In sentiment analysis, the number of occurrences of certain words (e.g. "good", "bad") is very important for distinguishing classes. MultinomialNB utilizes this frequency information effectively. In contrast, GaussianNB assumes continuous data and normal distribution, which is not suitable for text data. BernoulliNB only looks at the presence of (binary) words, thus ignoring their frequency. Therefore, MultinomialNB best aligns with the characteristics of text data and yields the best performance.

why Gaussian performs poorly, because Gaussian Naïve Bayes (GaussianNB) works well for normally distributed continuous data, such as temperature or height. wance w However, in text classification, features such as word frequency are discrete, non-negative and often skewed. This violates the assumptions of GaussianNB, making it difficult for the model to calculate probabilities accurately. As a result, it performs poorly on text data. In contrast, MultinomialNB is more suitable as it is designed to handle word frequency data directly.

4. Conclusions and Future Works

This study concludes that among the three Naïve Bayes algorithm variants—GaussianNB, BernoulliNB, and MultinomialNB—the Multinomial Naïve Bayes algorithm delivers the best performance for sentiment analysis of user reviews on the Mobile Legends: Bang Bang game. With an accuracy of 75%**,** it outperforms BernoulliNB (74%) and GaussianNB (50%). MultinomialNB is particularly effective because it aligns well with word frequency-based data representations commonly found in textual reviews. Meanwhile, GaussianNB performs poorly due to its assumption of continuous, normally distributed data, which is unsuitable for text, and BernoulliNB, while balanced, does not utilize word frequency information effectively. These findings affirm that choosing a model aligned with data characteristics is essential for accurate sentiment classification.

Future research can explore several directions to enhance sentiment analysis performance. One improvement is to implement more advanced classification methods such as Support Vector Machine (SVM), deep learning models like LSTM or BERT, or ensemble models (e.g., Random Forest, XGBoost) that can provide better generalization and handle more complex patterns in the data. Additionally, researchers may expand the study into multi-class sentiment analysis (e.g., positive, negative, neutral) for more granular insights, ensuring the neutral class is well-defined to handle ambiguous sentiments. Combining lexicon-based and supervised learning approaches, as well as applying feature selection techniques, can also improve classification performance and computational efficiency. Furthermore, increasing dataset diversity by including reviews from multiple platforms like Twitter or online forums can help improve the robustness and generalizability of the model. Lastly, incorporating sarcasm detection models or using contextual information can help better capture nuanced sentiments such as sarcasm, which often misleads conventional sentiment classifiers.

5. References.

[1] P. F. Lasamahu, D. Manongga, A. Iriani,"Sentiment analysis of online game users mobile legends with naïve bayes,"jan.2021, doi: https://repository.uksw.edu/handle/123456789/22401

[2] N. Khesya," Get to know flowcharts and pseudocodes in algorithms and programming", vol 1 des 2021 doi; https://doi.org/10.31219/osf.io/dq45e

[3] D. Normawati, S. A. Prayogi, "Implementation of Naïve Bayes classifier and confusion matrix on text-based sentiment analysis of twitter tada," Journal of Computer Science & Informatics (J-SAKTI) Volume 5 Number 2, September 2021, pp. 697-711 ISSN: 2548-9771/EISSN: 2549-7200,him:http://dx.doi.org/10.30645/j-sakti.v5i2.369

[4] A. Yusnita, S. Lailiyah, K. Saumahudi, "Application of the naïve bayes algorithm for the acceptance of new education participants," WICIDA informatics Vol.10 No.1(2021): January 2021,doi: https://doi.org/10.46984/inf-wcd.1194

[5] M. Y. Putra and D. I. Putri," Utilization of Naïve Bayes Algorithm and K-Nearest Neighbor for Classification of Majors of Grade XI Students," Jurnal TEKNO KOMPAK, Vol. 16, No. 2, P-ISSN: 1412-9663, E-ISSN: 2656-3525, Pages 176-187, 2022, doi: https://doi.org/10.33365/jtk.v16i2.2002

[6] N. Agustina,D. H. Citra, W. Purnama, C. Nisa, A. R. Kurnia," The Implementation of Naïve Bayes Algorithm for Sentiment Analysis of Shopee Reviews on Google Play Store," Indonesian Research and Publication Institute (IRPI) MALCOM:, Indonesian Journal of Machine Learning and Computer Science Journal Homepage: Vol. 2, Iss. 1 April 2022, pp: 47-54P-ISSN: 2797-2313 ISSN(E): 2775-8575, doi: https://journal.irpi.or.id/index.php/malcom

[7] A. I. Tanggraeni, M. N. N. Sitokdana, "Sentiment Analysis of E-Government Applications on Google Play Using the Naïve Bayes Algorithm," Journal of Informatics Engineering and Information Systems ISSN 2407-4322 Vol. 9, No. 2, June 2022, Pages 785-795 E-ISSN 2503-2933

[8] M. N. Rahman, "Performance analysis of the use of stopwords and stemming in sentiment analysis with the naive bayes classification approach,"Institutional Repository of Syarif Hidayatullah State Islamic University,2-Nov-2022 doi: https://repository.uinjkt.ac.id/dspace/handle/123456789/65210

[9] S. A. Rizaldi, S. Alam, I. Kurniawan, "Sentiment analysis of JMO (Jamsostek Mobile) application users on the Google Play Store using the naive bayes method," STORAGE - Scientific Journal of Engineering and Computer Science, Vol. 2No. 3, August 2023,109-117e-ISSN: 2828-5344, doi: 10.55123

[10] C. G. Indrayanto, D. E. Ratnawati2, B. Rahayudi, "Sentiment analysis of user review data of mypertaminadi indonesia application on google play store using random forest method," Journal of Information Technology and Computer Science Development-ISSN: 2548-964XVol. 7, No. 3, March 2023, pp. 1131-1139, doi: http://j-ptiik.ub.ac.id

[11] D. Rifaldi, A. Fadlil, Herman, "Preprocessing Techniques in Text Mining Using Mrntal Health Tweet Data," DECODE: Journal of Information Technology Education ISSN Print: 2775-2984 | ISSN Online: 2775-1813 Vol. 3, No. 2, September 2023, Pages: 161-171, Doi: http://dx.doi.org/10.51454/decode.v3i2.131

[12] K. D. Indarwati, H. Februaryyanti,Sentiment Analysis of Gojek Application Service Quality UsingKDNaive Bayes Classifier Method",JATISI, March 2023**,** Volume: 10 Number: 1 doi: https://doi.org/10.35957/jatisi.v10i1.2643

[13] J. .A. Rieuwpassa, Sugito, T. Widiharih, "Implementation of the naïve bayes classifier method for classifying sentiments of Netflix user reviews on Google Play," GAUSSIAN JOURNAL, Volume 12, Number 3, Year 2023, Pages 362-371, ISSN: 2339-2541, doi: https://ejournal3.undip.ac.id/index.php/gaussian/

[14] S. Bayulianto, I. Purnamasari, M. Jajuli,"Prediction of the winning rate of the Mobile Legends Professional League Indonesia Season 9 using the Naive Bayes algorithm",JIPI (Scientific Journal of Informatics Research and Learning), Vol. 8, No. 2, June 2023, Pp. 538-55,**doi:** https://doi.org/10.29100/jipi.v8i2.3562

[15] D. Wijaya, R. A. Saputra, F. Irwiensyah, "Sentiment Analysis of National Digital Samsat Application Reviews on Google Playstore Using the Naïve Bayes Algorithm," KLIK: Scientific Study of Informatics

and Computers ISSN 2723-3898 (Online Media) Vol 4, No 4, February 2024, Pages 2369-2380 DOI 10.30865/klik.v4i4.1738

[16] T. R. P. Hermawan, A. R. Dzikrillah, "Application of Naïve Bayes Method for Sentiment Analysis on ChatGPT Application User Reviews on Google Play Store." Building of Informatics, Technology and Science (BITS) Volume 6, No 1, June 2024 Page: 430-439 ISSN 2684-8910 (print media) ISSN 2685-3310 (online media) DOI 10.47065/bits.v6i1.5400

[17] A. D. A. S. Surya, Y. Yamasari, “Application of Naïve Bayes (NB) Algorithm for Heart Disease Classification,” JINACS: Volume 05 Number 03, 2024 (Journal of Informatics and Computer Science), ISSN: 2686-2220, doi: 10.26740/jinacs.v5n03.p447-455

[18] I. B. Setiawan, J. Maulindar, Nurchim, “Application of naïve bayes algorithm for sentiment analysis in digital health applications,” G-Tech: Jurnal Teknologi Terapan Vol. 8, No. 4, October 2024, pp. 2301-2312 E-ISSN: 2623-064X | P-ISSN: 2580-8737,doi: https://doi.org/10.70609/gtech.v8i4.5020

[19] A. F. N. Aziza, V. P. Ramadhan, “Comparison of Naïve Bayes and K-NN in Sentiment Analysis on Twitter Regarding the Victory of Candidate Pair 02, vol 12 no 2 2024 J-INTECH(journal of information and technology, doi: https://doi.org/10.32664/j-intech.v12i02.1305

[20] N. Z. B. Jannah, Kusnawi,” Comparison of Naïve Bayes and SVM in Sentiment Analysis of Product Reviews on Marketplaces, vol 8 no 2 2024 Sinkron (jurnal& penelitian Informatika, DOI: 10.33395/sinkron.v8i2.13559

[21] R. I. Agustin,” Comparison of naïve bayes and SVM Algorithms Umyuk Twitter Sentiment Analysis of Corruption of Rice Social Assistance during Pandemic Period’’,vol 12 no.2 2024 doi: https://doi.org/10.23960/jitet.v12i2.4020

[22] H. Firda, P. Putra, N. R. Oktadini, P. E, Sevtiyuni,” Comparison of Rating-based and Inset Lexicon-based Labeling in Sentiment Analysis Using SVM (Case Study: GoBiz Application Review on Google Play Store), Sistemasi: Jurnal Sistem Informasi Volume 14, Number 2, 2025: 516-528 ISSN:2302-8149 e-ISSN:2540-9719 doi: http://repository.unsri.ac.id/id/eprint/166015

[23] D. Ramadhani, D. Fadel, Bintang A. Nugroho, "User Sentiment Analysis Towards the Whatsapp Application on the Google Play Platform Using the Natural Language Processing Method," Journal of Computer Science (JIK) Volume 8, Number 1, March 2025: 39-49 E-ISSN: 2746-7163,

[24] M. Z. Siregar, A. M. Elhanafi, D. Irwan, "Sentiment analysis of social media application reviews on Google Play using the naïve bayes algorithm," JATI (Informatics Engineering Student Journal) Vol. 9 No. 2, April 2025, doi: https://doi.org/10.36040/jati.v9i2.12841