JUSIKOM PRIMA (Jurnal Sistem Informasi dan Ilmu Komputer Prima)
Vol. 7 No. 1, August 2023 E-ISSN : 2580-2879

SENTIMENT ANALYSIS OF MYPERTAMINA APPLICATION USING SUPPORT VECTOR MACHINE AND NAÏVE BAYES ALGORITHMS

Ongki Sopie Simbolon*[*1], Murni Esterlita Manullang[2], Stevin Alvarez[3], Lolo Frans M. Brutu[4], Evta Indra[5]
[1,2,,3,4,5]Universitas Prima Indonesia
Jl.Sampul No.3, Sei Putih Bar., Kec. Medan Petisah, Medan City, North Sumatra 20118
Email : onkysimbolon@gmail.com*

ABSTRACT- In line with the needs of the community and the progress of the times in the advanced field of fintech, cash payments are currently considered insecure as well as ineffective and efficient. To run a non-cash or cashless transaction program presently run by the government, PT. Pertamina invites the public to use E-Payment from the My Pertamina application in collaboration with LinkAja. In this study, the sentiments of MyPertamina application users will be analyzed based on reviews on the Google Play Store. Review data will be analyzed to determine whether the check has positive, negative, or neutral sentiments. The data analysis stage is text preprocessing to change uppercase to lowercase, clearing text, separating text, taking important words, changing essential phrases, and labeling data into positive, negative, and neutral classes. As well as the classification and evaluation of results. This study used the Support Vector Machine (SVM) and Naïve Bayes classification methods. To evaluate the results, the confusion matrix was used to test the accuracy, Precision, recall, and F1 score value. The classification results obtained the highest accuracy value for the Support Vector Machine (SVM) method, which had accuracy (68.50%), precision (70.00%), recall (69.70%), and F1 score (68.46%). Meanwhile, the Naïve Bayes method has performance with accuracy (63.00%), precision (63.90%), recall (61.34%), and F1 score (59.55%).

KEYWORDS: Classification, Naïve Bayes, Review, Sentiment, Support Vector Machine

1. INTRODUCTION

The presence of non-cash transactions where Bank Indonesia has legally issued non-cash payments, which as a transaction method is considered more protected, adequate, and efficient compared to cash transactions (Muhammad et al. Moeliono, 2020). At a time when technology is increasingly advanced, payment instruments have switched to non-cash payments and electronic payments. The increase in users and electronic transactions aligns with the rise in the total money in the environment. To run a non-cash or cashless transaction program currently run by the government, PT. Pertamina invites the public to use E-Payment from the My Pertamina application in collaboration with LinkAja. Even though gas stations that can serve transactions through MyPertamina are widely spread in various cities and districts, MyPertamina is not very well known among the public (Maria et al., 2023). Research is needed to determine how much people respond after using E-Payment from the MyPertamina application to determine whether the MyPertamina application meets user standards for the community.

Many previous studies have used Machine Learning to determine sentiment analysis. One of them is research (Kelvin et al., 2022) "Comparative Analysis of Corona Virus Disease-2019 (Covid19) Sentiments on Twitter Using Logistic Regression and Support Vector Machine (SVM)

Methods". This research classified positive, negative, or neutral sentiments toward the Covid-19 case on Twitter. From this research, the results of the accuracy of the two methods were that the Support Vector Machine (SVM) method produced the highest accuracy value of 92.13% in training data and 91.15% in test data. In contrast, the Logistic Regression method obtained an accuracy value of 87.79% in training data and 87.68% in test data. From the accuracy value obtained, the final result of this research is that the Support Vector Machine (SVM) method excels in classifying Twitter sentiment about the Covid-19 case.

Support Vector Machine (SVM) and Naïve Bayes are relatively easy to implement and have extensive support from various libraries and frameworks in Python, such as Scikit-learn and NLTK. And also has a relatively easy interpretation, so it can help understand the factors that influence the sentiment in the text. Based on the problems described in the previous paragraph, the authors suggest creating a Machine Learning model using the Support Vector Machine (SVM) and Naïve Bayes algorithms to predict the sentiment labels of MyPertamina application users, whether they are positive, negative, or neutral. So that PT. Pertamina itself can improve the MyPertamina application so that it is more comfortable for users to use. The parameters used in creating the model are username, date, rating, and user reviews of the MyPertamina application obtained from the Google Play Store.

2. RESEARCH METHODS

The first step in this research methodology is to collect MyPertamina application review data on the Play Store. The next step is to label the data set using the Vader Lexicon dictionary with positive, negative, and neutral labels. After the labeling procedure is completed, data preprocessing involves case folding, cleaning, tokenizing, stemming, and filtering. Machine learning, especially Support Vector Machine and Naive Bayes, will be used for processing these features. Here's the progress of his research.

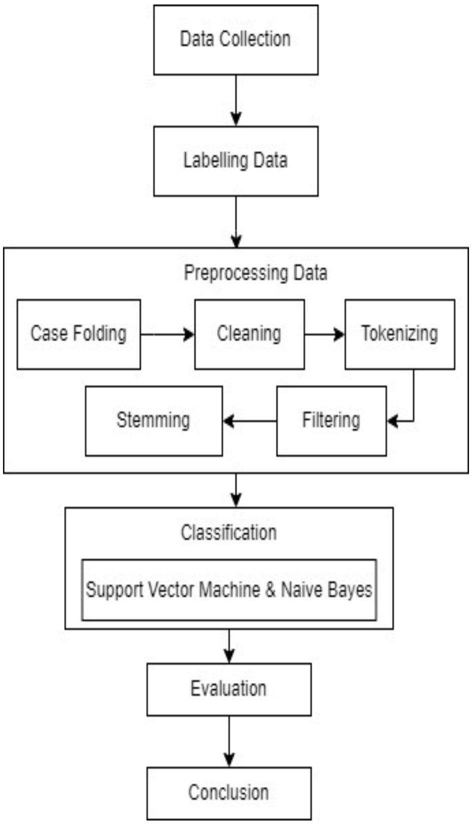


Figure 1. Research Flow
(Source: Personal Documentation, 2023)

2.1 Data Collection

Data was collected by retrieving user review data from the MyPertamina application from the comments column on the Google Play Store using the Python Scrapper library, which runs on Google Collaboratory with the Python programming language.

	reviewId	userName	userImage	content	score	thumbsUpCount	reviewCreatedVersion	at	replyContent	repliedAt
0	f4a31a05-7cbd-41ba-a752-791177cbf17d	31_Nunik Hariadi Putri	https://play-lh.googleusercontent.com/a/AGNmyx...	Flow aplikasinya tidak jelas, membingungkan pe...	1	297	3.7.4	2023-02-05 02:54:38	None	NaT
1	44d46e2f-c5d0-482b-87b0-2b32f5d8dcdb	Ahmad Tantowi dastan	https://play-lh.googleusercontent.com/a-/ACB-R...	Kecewa sekaligus membingungkan Setiap mengedit...	1	175	3.7.4	2023-02-09 03:26:35	None	NaT
2	4abbf9b-357e-49ca-b581-5c001b1b16cb	A Zi	https://play-lh.googleusercontent.com/a-/ACB-R...	Aplikasi Tergak jelas yg pernah ada di play s...	1	136	3.7.5	2023-02-13 11:03:36	None	NaT
3	b4d1e348-84e9-4696-a342-4c7652d5828f	Gebi Chikita	https://play-lh.googleusercontent.com/a/AGNmyx...	Bikin aplikasi cuma bikin tambah susah masyara...	1	33	3.7.5	2023-02-21 12:39:48	None	NaT
4	6e861d25-99e7-47aa-9550-d7d0d8fa774f	Sekar Cel Korlek	https://play-lh.googleusercontent.com/a-/ACB-R...	Mengambil gambar di aplikasi tidak bisa fokus...	1	67	3.7.5	2023-02-14 00:19:45	None	NaT
5	73f5dc52-906d-48d4-b2c4-e1b88d8b628f	Albert Lim	https://play-lh.googleusercontent.com/a-/ACB-R...	Aplikasi bikin bingung Pembayaran juga hanya b...	1	6	3.7.5	2023-02-23 02:03:38	None	NaT
6	f9f6d021-3084-4885-8004-b9cda7020c1f	epick Aich	https://play-lh.googleusercontent.com/a/AGNmyx...	Aplikasinya banyak Bug. Login berulang suruh m...	1	12	3.7.5	2023-02-18 08:49:42	None	NaT
7	fdb5d501-b332-4ae8-a8fb-d0ebab405ff9	Su Hendrik Thung	https://play-lh.googleusercontent.com/a-/ACB-R...	Aplikasi suka macet, pas mau update data layar...	1	1	3.7.5	2023-02-26 06:28:24	None	NaT
8	2834d930-d342-4705-a338-efa5d3d8f190	Yus Permadi	https://play-lh.googleusercontent.com/a-/ACB-R...	Sangat membingungkan, ketika mw menambahkan dat...	1	38	3.7.5	2023-02-12 01:13:50	None	NaT
9	79a4a785-14a4-48e0-aa73-3cad65037142	K Kurniawan	https://play-lh.googleusercontent.com/a/AGNmyx...	Ini itu aplikasi aneh.. Disuruh foto STNK, sud...	1	19	3.7.4	2023-02-07 07:13:47	None	NaT

Figure 2. Scraping Result Data
(Source: Personal Documentation, 2023)

2.2 Labelling Data

Data labeling in this study was assisted by the Vader Lexicon library or dictionary (lexicon). This model utilizes a dictionary of words that have been given a positive, negative, or neutral sentiment score and linguistic rules to calculate the overall sentiment score of a text.

2.3 Preprocessing Data

At this stage, it is done by using a library in the Python programming language. Data preprocessing is performed by Case Folding, Cleaning, Tokenizing, Filtering, and Stemming stages to produce clean and ready-to-use data processed later. The process carried out is as follows.

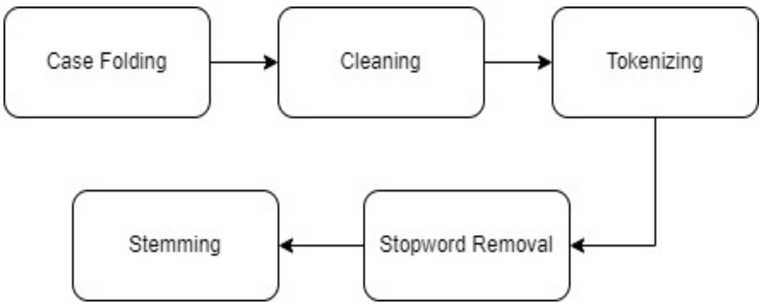


Figure 3. Preprocessing Stage
(Source: Personal Documentation, 2023)

2.4 Classification

Before classifying with Support Vector Machine (SVM) and Naïve Bayes, the review data will be divided into train and test data. At this stage, the training data & test data will be classified into Support Vector Machine (SVM) and Naïve Bayes classifications using the Python programming language at the Google Collaboratory. The classification output results from the accuracy, Precision, recall, and f1 score of each classification that will be used.

2.5 Evaluation

The results evaluation stage will be assisted by the confusion matrix method. Confusion Matrix is an n x n matrix used to evaluate a classification model's performance, where n is the number of class targets. The matrix contains the numbers from the actual value with the predicted value generated from the classification model to find out how well the performance of the classification model is (Bhanujyothi H C, Dr.Chetana Tukkoji, 2021).

a. Accuracy shows how accurately the model correctly classifies the data.

$$\text{accuracy} = \frac{(TP+TN)}{(\text{TP+FP+FP+FN})} \text{ x } 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

b. Precision describes the accuracy of the prediction results given by the model.

$$\text{precision} = \frac{(TP)}{(\text{TP+FP})} \text{ x } 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (2)$$

c. Recall describes the success of the model in retrieving information.

$$\text{recall} = \frac{(TP)}{(\text{TP+FN})} \text{ x } 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (3)$$

d. The F1 score is obtained by combining precision and recall values.

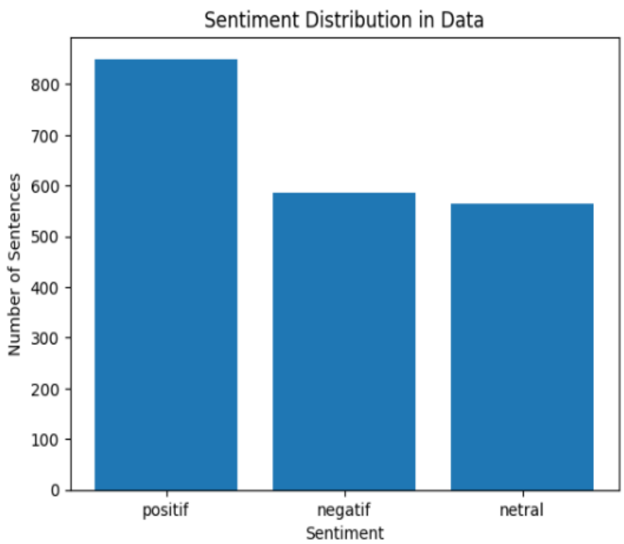
$$\text{f1 score} = 2 \text{ x } \frac{(recall\ x\ precision)}{(recall\ x\ precision)} \text{ x } 100\% \quad \ldots\ldots\ldots\ldots\ldots(4)$$

3. RESULTS AND DISCUSSION

3.1 Labelling Data

This stage aims to classify review or review data into positive, negative, and neutral sentiment classes. In this study, data labeling was assisted using the Python library, namely VADER (Valence et al.). The library is a sentiment analysis algorithm that utilizes a lexicon. The determination of positive, neutral, and negative classes with Vader Lexicon is based on polarity values and composite slots, a combination of the three. Calculating the polarity value is said to be positive if the composite value is more than or equal to 0.05. It is negative if the composite value is less than or equal to -0.05 and neutral if it is in the middle or 0.

Figure 4. Data Labelling Results



(Source: Personal Documentation, 2023)

Figure 4 can be seen as a visualization of the results of data labeling using the Vader Lexicon. It can be seen in the picture that most reviews on the MyPertamina application are positive, with a total of 850 reviews. And for negative and neutral values, it is 585 for negative reviews and 565 for neutral reviews. This shows that the MyPertamina application has received a fairly good response from users.

3.2 Preprocessing Data

a. Case Folding

Case Folding is a text preprocessing stage that aims to change all upper case letters in the document into lower case letters.

Table 1. Result of Case Folding

Before	After
Aplikasinya masih menampilkan popup even promo, auto log off. SPBU yg bs trx pake aplikasi bnyk yg ga aktifin aplikasinya dgn alas an system/aplikasi gangguan. Aneh banget	aplikasinya masih menampilkan popup even promo, auto log off. spbu yg bs trx pake aplikasi bnyk yg ga aktifin aplikasinya dgn alas an system/aplikasi gangguan. aneh banget

b. Cleaning

Cleaning is a text preprocessing stage that aims to clean text from tabs, new lines, back slices, mentions, links, hashtags, and URLs. This process begins by removing delimiters, namely symbols and punctuation marks in the text, such as @, $, &, full stop (.), comma (,), question mark (?), and exclamation point (!).

Tabel 2. Result of Cleaning

Before	After
subsidi tepat dari mana ?? yang beli ketengan tetap bisa tanpa qr code yang beli untuk kendaraan sendiri malah di susashin hebat memang ini Indonesia!! #mypertamina	aplikasinya masih menampilkan popup even promo, auto log off. spbu yg bs trx pake aplikasi bnyk yg ga aktifin aplikasinya dgn alas an system/aplikasi gangguan. aneh banget

c. Tokenizing

Tokenizing is a text preprocessing stage that aims to engrave text into words called tokens. The purpose of tokenizing is so that data can be processed at a later stage, namely removing the extended stopword (filtering).

Tabel 3. Result of Tokenizing

Before	After
tidak bisa login nik dan password sudah dimasukin .bilang salah .gimana bisa login mau rubah data tidak bisa.garis 3 kedap kedip lanjut nga bisa kembali nga bisa susah sekali.	‘tidak’, ‘bisa’, ‘login’, ‘nik’, ‘dan’, ‘password’, ‘sudah’, ‘dimasukin’, ‘.bilang’, ‘salah’, ‘.gimana’, ‘bisa’, ‘login’, ‘mau’, ‘rubah’, ‘data’, ‘tidak’, ‘bisa’, ‘garis’, ‘3’, ‘kedap’, ‘kedip’, ‘lanjut’, ‘nga’, ‘bisa’, ‘kembali’, ‘nga’, ‘bisa’, ‘susah’, ‘sekali’,

d. Filtering

Filtering is the text processing stage by taking important words from the token results using a stoplist algorithm (removing less important words) or wordlist (saving important words).

Tabel 4. Result of Filtering

Before	After
‘aplikasi’, ‘gak’, ‘akurat’, ‘semua’, ‘semua’, ‘dibuat’, ‘sulit’	‘aplikasi’, ‘akurat’, ‘sulit’

e. Stemming

Stemming is the stage of text processing to get the base word from a word that has been affixed with the assumption that these words have the same meaning and significance.

Table 5. Result of Stemming

Before	After
‘pemerintah’, ‘gak’, ‘jelas’, ‘buat’, ‘aplikasi’, ‘terniat’, ‘nya’, ‘untuk’, ‘ngebantu’, ‘masyarakat’, ‘malah’, ‘nyusahin’	‘niat’, ‘bantu’, ‘susah’

3.3 Classification

Classification with Support Vector Machine and Naïve Bayes is obtained by testing based on the training data and test data to be tested. The training data and test data tested are 80% training data & 20% test data. Using these two data, accuracy, Precision, recall, and f1-score values will be tested from classification with Support Vector Machine (SVM) and Naive Bayes. After processing, all classification of training data and test data with Support Vector Machine

(SVM) will be compared with the Naïve Bayes method.

Table 6. Result of Classification

Method	Score			
	Accuracy	Precision	Recall	F1 Score
Support Vector Machine	68.50%	70.00%	69.70%	68.46%
Naïve Bayes	63.00%	63.90%	61.34%	59.55%

The table shows the performance evaluation results of the two classification models: Support Vector Machine (SVM) and Naïve Bayes. Four performance evaluation metrics are calculated: Accuracy, Precision, recall, and F1 score. Accuracy measures how many cases the model correctly classifies. In the table, SVM has an accuracy of 68.50%, while Naïve Bayes has an accuracy of 63.00%. This means that SVM is more accurate in classifying data than Naïve Bayes. Precision measures how many positive results are positive. In the table, SVM has a precision of 70.00%, while Naïve Bayes has a precision of 63.90%. This means that SVM is better at identifying positive cases than Naïve Bayes. Recall measures how many positive cases the model has identified. In the table, SVM has a recall of 69.70%, while Naïve Bayes has a recall of 61.34%. This means that SVM is better at identifying positive cases than Naïve Bayes. The F1 score combines Precision and recall to provide a more representative value of the overall model performance. In the table, SVM has an F1 score of 68.46%, while Naïve Bayes has an F1 score of 59.55%.

3.4 Evaluation

1. Confusion Matrix from the Support Vector Machine method

Based on the Confusion Matrix, how much data is predicted into the correct class can be described. The magnitude of the classification data in the negative category that can be predicted correctly into the negative class is called a true negative. True neutral is the amount of neutral category classification data that can be accurately predicted into a neutral class. The magnitude of the classification data in the positive category that can be predicted correctly into the positive class is called true positive. The number of observational data is in a positive category, but there are prediction errors called false positives. The number of observational data is in the neutral category, but there is a false neutral prediction error. The number of observational data is in the negative category, but there are prediction errors called false negatives.

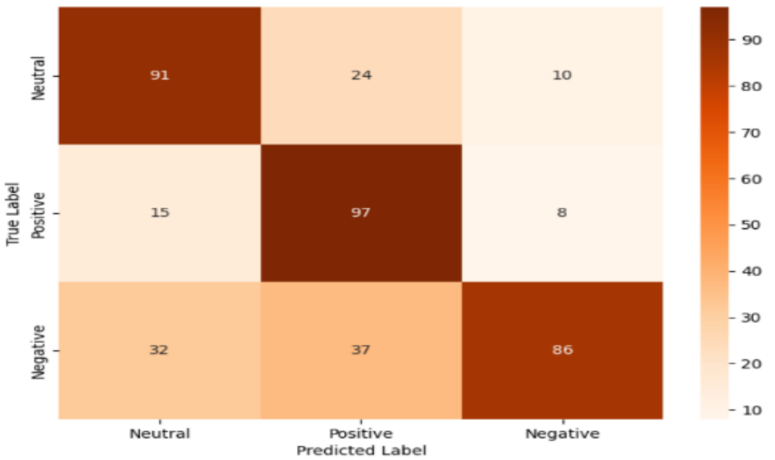


Figure 5. Confusion Matrix model Support Vector Machine (SVM)
(Source: Personal Documentation, 2023)

Table 7. Classification of TP, TN, FP, and FN values (Support Vector Machine)

Model Predictions	True Positive	True Negative	False Positive	False Negative
Positive	91	228	47	34
Neutral	97	219	61	23
Negative	86	227	18	69

By looking at the grouping table of TP, TN, FP, and FN values, accuracy, Precision, recall, and f1-score values can be calculated from the Support Vector Machine (SVM) classification model:

a. Accuracy

$$\text{Accuracy} = \frac{(91+97+86)}{(91 + 24 + 10 + 15 + 97 + 8 + 32 + 37 + 86)} \text{ x } 100\% = 68.5\% \text{ …………………(5)}$$

Based on the above calculations, it is known that the results of sentiment classification with test data provide an overall accuracy of 68.5%

b. Precision

$$\text{P (Positif)} = \frac{(97)}{(97+24+37)} = 0.668 \text{ ……………………………………………………(6)}$$

$$\text{P (Negatif)} = \frac{(86)}{(86+8+10)} = 0.866 \text{ ……………………………………………………(7)}$$

$$\text{P (Netral)} = \frac{(91)}{(91+15+32)} = 0.688 \text{ ……………………………………………………(8)}$$

$$\text{Precision} = \frac{P(positif)+P(negatif)+P(netral)}{total\ kelas} \text{ x } 100\% \text{ ……………………......………..(9)}$$

$$= \frac{0.668+0.866+0.688}{3} * 100\% = 70.0\% \text{ …………………….……………………(10)}$$

The average precision value of the overall precision value for each sentiment class is 70%.

c. Recall

$$\text{R (Positif)} = \frac{(97)}{(15+97+8)} = 0.811 \text{ …………………………………………………(11)}$$

$$\text{R (Negatif)} = \frac{(86)}{(32+37+86)} = 0.594 \text{ ……………………………………………...(12)}$$

$$\text{R (Netral)} = \frac{(91)}{(91+24+10)} = 0.714 \text{ …………………………………………………(13)}$$

$$\text{Recall} = \frac{R(positif)+R(negatif)+R(netral)}{total\ kelas} \text{ x } 100\% \text{ ………………………………..(14)}$$

$$= \frac{0.811+0.594+0.714}{3} * 100\% = 69.7\% \text{ ………………..…………………..(15)}$$

The average recall value of the overall precision value for each sentiment class is 69.7%.

d. F1 Score

F1 Score (Positif) = 2 x $\frac{(0.688 \; x \; 0.811)}{(0.688+0.811)}$ = 0.705 ....(16)

F1 Score (Negatif) = 2 x $\frac{(0.866 \; x \; 0.594)}{(0.866+0.594)}$ = 0.86 ....(17)

F1 Score (Netral) = 2 x $\frac{(0.688 \; x \; 0.714)}{(0.688+0.714)}$ = 0.700 ....(18)

F1 Score = $\frac{F1\;Score(positif) + F1\;Score(negatif) + F1\;Score(netral)}{total\;kelas}$ x 100% ....(19)

$= \frac{0.700+0.733+0.705}{3} * 100\% = 68.5\ \%$ ....(20)

Based on the calculations, the average recall value of the overall f1 score for each sentiment class is 68.5%

2. Confusion Matrix from the Naïve Bayes method

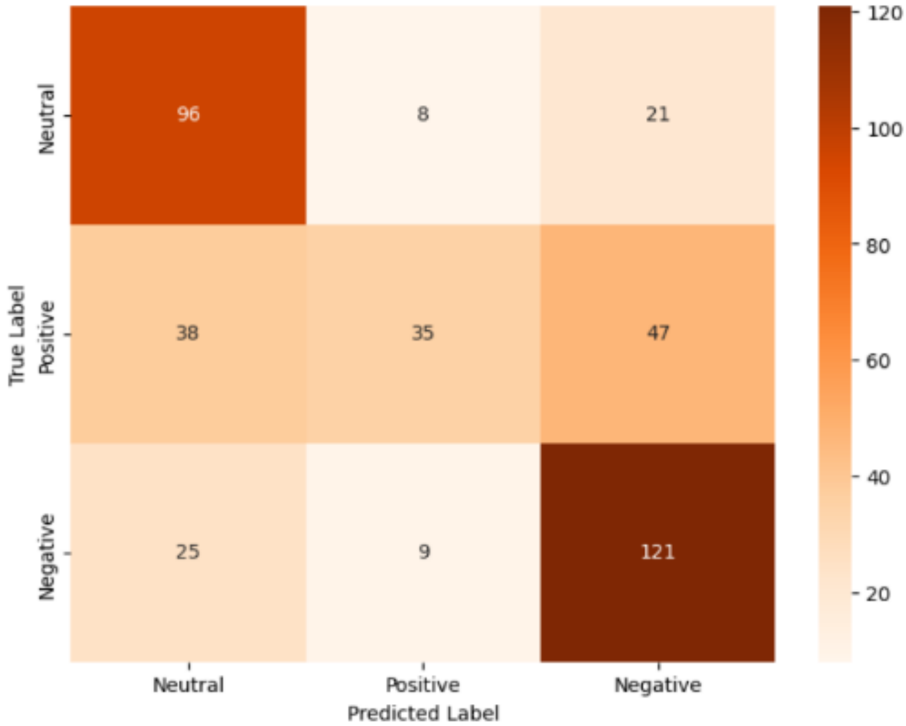


Figure 6. Confusion Matrix model Naïve Bayes
(Source: Personal Documentation, 2023)

The Confusion Matrix of the Naïve Bayes model is a table that shows the predicted results of the Naïve Bayes model on test data and train data. The confusion matrix can be used to test the results of classification values in models such as accuracy, Precision, recall, and F1 scores. As before in the Support Vector Machine (SVM) model, the confusion matrix in the Naïve Bayes model also consists of four cells, namely True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN).

Table 8. Classification of TP, TN, FP, and FN values (Naïve Bayes

Model Predictions	True Positive	True Negative	False Positive	False Negative
Positive	96	212	63	29
Neutral	35	263	17	85
Negative	121	177	68	34

By looking at the grouping table of TP, TN, FP, and FN values, accuracy, Precision, recall, and f1-score values can be calculated from the Naïve Bayes classification model:

a. **Accuracy**

$$\text{Accuracy} = \frac{(96+35+121)}{(96 + 8 + 21 + 38 + 35 + 47 + 25 + 9 + 121)} \text{ x } 100\% = 63\ \% \quad \text{……………………..(21)}$$

Based on the calculations, it is known that the results of sentiment classification with test data & train data (80%:20%) in this study provide an overall accuracy of 63%.

b. **Precision**

$$\text{P (Positif)} = \frac{(35)}{(8+35+9)} = 0.661 \quad \text{………………………………………………………….(22)}$$

$$\text{P (Negatif)} = \frac{(121)}{(21+47+121)} = 0.655 \quad \text{……………………………………………………..(23)}$$

$$\text{P (Netral)} = \frac{(96)}{(96+38+25)} = 0.639 \quad \text{…………………………………………………………...(24)}$$

$$\text{Precision } = \frac{P(positif)+P(negatif)+P(netral)}{total\ kelas} \text{x } 100\ \% \quad \text{…………………………………..(25)}$$

$$= \frac{0.661+0.655+0.639}{3} * 100\% = 63.9\ \% \quad \text{……………………………………..(26)}$$

Based on calculations, the average precision value of the overall precision value for each sentiment class is 63.9%

c. **Recall**

$$\text{R (Positif)} = \frac{(35)}{(38+35+47)} = 0.296 \quad \text{……………………………………………………….(27)}$$

$$\text{R (Negatif)} = \frac{(121)}{(25+9+121)} = 0.830 \quad \text{……………………………………………………….(28)}$$

$$\text{R (Netral)} = \frac{(96)}{(96+8+21)} = 0.748 \quad \text{………………………………………………………….(29)}$$

$$\text{Recall} = \frac{R(positif)+R(negatif)+R(netral)}{total\ kelas} \text{x } 100\ \% \quad \text{……………………………………...(30)}$$

$$= \frac{0.296+0.830+0.748}{3} * 100\% = 61.3\ \% \quad \text{…………………………………………..(31)}$$

Based on the calculations, the average recall value of the overall precision value for each sentiment class is 61.3%

d. **F1 Score**

F1 Score (Positif) = 2 x $\frac{(0.661 \ x \ 0.296)}{(0.661+ \ 0.296)}$ = 0.406 ……………………………………..(32)

F1 Score (Negatif) = 2 x $\frac{(0.655 \ x \ 0.830)}{(0.655+ \ 0.830)}$ = 0.73 ……………………………………..(33)

F1 Score (Netral) = 2 x $\frac{(0.639 \ x \ 0.748)}{(0.639 \ + \ 0.748)}$ = 0.689 ……………………………………(34)

F1 Score = $\frac{F1 \ Score(positif) \ + \ F1 \ Score(negatif) \ + \ F1 \ Score(netral)}{total \ kelas}$ x 100 ……………...(35)

= $\frac{0.406+0.732+0.689}{3} * 100\% = 60.9$ % ……………………………………(36)

Based on the calculations, the average recall value of the overall f1 score for each sentiment class is 60.9%

CONCLUSION

Based on a comparison of the test data and train data used from the performance results of the Support Vector Machine (SVM) classification model and the Naïve Bayes classification, it can be concluded that the Support Vector Machine (SVM) model has a better performance compared to the Naïve Bayes model. This can be seen from the higher accuracy (68.50%), Precision (70.00%), recall (69.70%), and F1 score (68.46%) in the Support Vector Machine (SVM) model compared to the Naïve Bayes model with accuracy (63.00%), precision (63.90%), recall (61.34%), and F1 score (59.55%). Therefore, the Support Vector Machine (SVM) model can be chosen as a more optimal model for classifying test and training data. Of the 2,000 review samples, it was found that most of the MyPertamina application review data were negative review data, with class data of 585 negative reviews, 850 positive reviews, and 565 neutral reviews. Based on the results of the output above, the application itself has received a fairly good response from the users of the application itself. It needs to be upgraded again for adjustments according to user needs.

BIBLIOGRAPHY

Ahmadi, M. I., Apriani, F., Kurniasari, M., Handayani, S., & Gustian, D. (2020). Sentiment Analysis Online Shop on the Play Store Using Method Support Vector Machine (Svm. Seminar Nasional …, 2020(Semnasif), 196–203. http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4101

Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. (2020). Sentiment Analysis on E-Sports for Education Curriculum Using Naive Bayes and Support Vector Machine. Jurnal Ilmu Komputer Dan Informasi, 13(2), 109–122. https://doi.org/10.21609/jiki.v13i2.885

Bhanujyothi H C, Dr.Chetana Tukkoji, V. J. S. T. J. S. B. (2021). Prognosis of Diabetes Mellitus using Machine Learning Techniques. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(5), 836–841. https://doi.org/10.17762/turcomat.v12i5.1491

Bourequat, W., & Mourad, H. (2021). Sentiment Analysis Approach for Analyzing iPhone Release using Support Vector Machine. International Journal of Advances in Data and Information Systems, 2(1), 36–44. https://doi.org/10.25008/ijadis.v2i1.1216

Damanik, F. J., & Setyohadi, D. B. (2021). Analysis of public sentiment about COVID-19 in Indonesia on Twitter using multinomial naive Bayes and support vector machine. IOP Conference Series: Earth and Environmental Science, 704(1). https://doi.org/10.1088/1755-1315/704/1/012027

Deta Kirana, Y., & Al Faraby, S. (2021). Sentiment Analysis of Beauty Product Reviews Using the K-Nearest Neighbor (KNN) and TF-IDF Methods with Chi-Square Feature Selection. Open Access J Data Sci Appl, 4(1), 31–042. https://doi.org/10.34818/JDSA.2021.4.71

Fransiska, S., & Irham Gufroni, A. (2020). Sentiment Analysis Provider by. U on Google Play Store Reviews with TF-IDF and Support Vector Machine (SVM) Method. Scientific Journal of Informatics, 7(2), 2407–7658. http://journal.unnes.ac.id/nju/index.php/sji

Hadju, S. F. N., & Jayadi, R. (2021). Sentiment analysis of Indonesian e-commerce product reviews using support vector machine-based term frequency inverse document frequency. Journal of Theoretical and Applied Information Technology, 99(17), 4316–4325.

Kaswidjanti, W., Himawan, H., & Silitonga, P. D. P. (2020). The accuracy comparison of social media sentiment analysis using lexicon based and support vector machine on souvenir recommendations. Test Engineering and Management, 82(3–4), 3953–3961.

Katta, P., & Hegde, N. P. (2019). A hybrid adaptive neuro-fuzzy interface and support vector machine based sentiment analysis on political Twitter data. International Journal of Intelligent Engineering and Systems, 12(1), 165–173. https://doi.org/10.22266/IJIES2019.0228.17

Kelvin, K., Banjarnahor, J., -, E. I., & NK Nababan, M. (2022). Analisis perbandingan sentimen Corona Virus Disease-2019 (Covid19) pada Twitter Menggunakan Metode Logistic Regression Dan Support Vector Machine (SVM). Jurnal Sistem Informasi Dan Ilmu Komputer Prima(JUSIKOM PRIMA), 5(2), 47–52. https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2365

Maria, R., Umayah, R. U., Mahardinny, S., Kalana, D. N., & Saputra, D. D. (2023). Analisis Sentimen Persepsi Masyarakat Terhadap Penggunaan Aplikasi My Pertamina Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier. 1, 1–10.

Muhammad Ibrahim, R., & Novandriani Karina Moeliono, N. (2020). Persepsi Konsumen Pada My Pertamina (Studi Pada Penggunaan My Pertamina Kota Bandung). Jurnal Ilmiah Mahasiswa Ekonomi Manajemen Accredited SINTA, 4(2), 396–413. http:jim.unsyiah.ac.id/ekm

Nurkholis, A., Alita, D., & Munandar, A. (2022). Comparison of Kernel Support Vector Machine Multi-Class in PPKM Sentiment Analysis on Twitter. Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 6(2), 227–233. https://doi.org/10.29207/resti.v6i2.3906

Padilah, B., Pratama, A. R., & Juwita, A. R. (2023). Analysis of Sentiment Adiraku App Reviews on Google Play Store Using Vector Machine Support Algorithm and Naïve Bayes. Jurnal Sisfotek Global, 13(1), 8. https://doi.org/10.38101/sisfotek.v13i1.2943

Pasaribu, J. S. (2020). Appilication of K-Means algorithm to predict consumer interest according to the season on place reservation and food online software. Journal of Physics: Conference Series, 1477(2). https://doi.org/10.1088/1742-6596/1477/2/022023

Prasetyo, I. R. M. R., Musthafa, A., & Taufiqurrahman. (2023). Comparison between naive bayes method and support vector machine in sentiment analysis of the relocation of the Indonesian capital. Jurnal Mantik, 7(2), 185–193.

Pratmanto, D., Rousyati, R., Wati, F. F., Widodo, A. E., Suleman, S., & Wijianto, R. (2020). App Review Sentiment Analysis Shopee Application in Google Play Store Using Naive Bayes Algorithm. Journal of Physics: Conference Series, 1641(1). https://doi.org/10.1088/1742-6596/1641/1/012043

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

https://doi.org/10.1088/1757-899x/1088/1/012045

Saddam, M. A., Dewantara, E. K., & Solichin, A. (2023). Sentiment Analysis of Flood Disaster Management in Jakarta on Twitter Using Support Vector Machines. Sinkron, 8(1), 470–479. https://doi.org/10.33395/sinkron.v8i1.12063

Sengkey, D. F., Jacobus, A., & Manoppo, F. J. (2019). Implementing Support Vector Machine Sentiment Analysis to Students' Opinion toward Lecturer in an Indonesian Public University. Journal of Sustainable Engineering: Proceedings Series, 1(2), 194–198. https://doi.org/10.35793/joseps.v1i2.27