RAUDHAH PUBLISHER

Journal of Social Science and Education Research
Volume 3, No. 1, 2026 - ISSN: 3066-6554
Website: https://jitir.seevalue.org/ojs33/index.php/jitir/index

Social Media Sentiment Analysis Using Machine Learning to Improve Digital Banking Services

Yulian Sani[1], Alex Winarno[2], Agus Maolana Hidayat[3]
Master of Business Administration Program, Faculty of Economics and Business, Telkom University, Bandung, Indonesia [1,2,3]

Keywords:
Sentimen Analysis
Digital Banking
Naïve Bayes
Super Vector Machine
Machine Learning
SWOT Analysis
Fishbone
5 Whys.

Abstract

The development of digital banking services has transformed the interaction patterns between consumers and banking products or services, including at PT Bank Digital Nusantara, which operates the ABC application as its primary mobile-based financial transaction platform. This study aims to analyze public sentiment toward the Bank's official social media content on Instagram and TikTok, as an effort to understand the perceptions, needs, and feedback, particularly from the bank's customers. Using 10,198 comments collected during the period from August 31, 2024, to August 31, 2025, this study applies two Machine Learning methods, namely Naïve Bayes and Support Vector Machine (SVM), to classify customer sentiment and compare the performance of both models. The sentiment analysis results are then integrated into a SWOT analysis to identify strengths, weaknesses, opportunities, and threats, which are further examined through Root Cause Analysis using the Fishbone Diagram and 5 Whys approach to systematically determine the underlying causes of issues. The study shows that leveraging sentiment analysis plays a strategic role in supporting data-driven decision-making, enhancing the company's responsiveness to service issues, strengthening brand awareness, and promoting the competitiveness of ABC by Nusantara as a digital bank in Indonesia. These findings are expected to contribute to the development of digital banking business strategies while enriching the academic literature on the application of sentiment analysis to improve service quality and customer experience.

Corresponding Author:
Yulian Sani
Email:
yuliansani@student.telkomuniversity.ac.id

INTRODUCTION

The development of digital technology has driven significant changes in how consumers interact with banking services. PT Bank Digital Nusantara is one of the digital banks in Indonesia that offers a branchless banking experience through the Abc application as its primary financial transaction platform. The Bank currently utilizes social media, particularly Instagram and TikTok, as channels to build communication with customers, obtain feedback, and

understand public perceptions of the services provided (Drus & Khalid, 2019). However, the absence of an adequate opinion-monitoring system may lead to uncontrolled issues, reputational risks, and a decline in customer trust.

In the financial domain, sentiment analysis has been applied to examine subjective expressions and their relationship with organizational performance, indicating its relevance for strategic evaluation in financial services (Zhong & Ren, 2022).

Sentiment analysis serves as a strategic approach to identifying positive, negative, and neutral opinions expressed by users (Axhiu et al., 2014; Liu, 2022; Pak & Paroubek, 2010; Pang & Lee, 2008). This study focuses on employing sentiment analysis on 10,198 comments collected from the Bank's social media platforms during the period of 31 August 2024 to 31 August 2025. The implementation of sentiment analysis not only enables early detection of complaints or emerging issues but also assists the Bank in understanding the preferences and needs of customers who actively use these visual and interactive platforms.

Previous studies have extensively discussed the evolution, application domains, and methodological trends of sentiment analysis across various online platforms. Systematic and tertiary reviews highlight that sentiment analysis has been widely applied to social media data to capture public opinion, user perception, and behavioral patterns, particularly in large-scale and data-intensive environments(Chen et al., 2017; Ligthart et al., 2021; Rodríguez-Ibánez et al., 2023).

Sentiment analysis has also been adopted across multiple domains, including education, business, and public services, demonstrating its flexibility as a methodological tool for extracting meaningful insights from unstructured textual data in diverse application contexts (Pooja & Bhalla, 2022; Pradhan & Merline, 2023).

Recent empirical studies demonstrate that sentiment analysis on social media effectively captures customer perception and public opinion across different linguistic and cultural contexts, using various machine learning approaches to evaluate service quality, social issues, and user satisfaction (Alsemaree et al., 2024; Fauzi et al., 2024).

Beyond service evaluation, sentiment analysis has also been utilized to detect anomalies, misinformation, and distorted narratives in social media environments, highlighting its broader analytical potential (Bhardwaj et al., 2024).

To produce a more comprehensive understanding, this research employs two Machine Learning methods: Naïve Bayes and Support Vector Machine (SVM) (Hastie et al., 2005; Joachims, 1998; McCallum, 1998; Medhat et al., 2014). Naïve Bayes offers advantages in fast and efficient classification processes, while SVM is known for delivering more accurate classification results through optimal data separation. The use of both methods allows for performance comparison and minimizes bias in sentiment interpretation, thereby providing a stronger and more reliable depiction of public perception.

The sentiment classification results are subsequently used as input for a SWOT analysis to identify strengths, weaknesses, opportunities, and threats relevant to the development of the Bank's services (David et al., 2017). Findings related to weaknesses are further examined through Root Cause Analysis (RCA) using the Fishbone Diagram and 5 Whys method to systematically trace the underlying causes of problems (Divya, 2019). The combination of these approaches provides deeper insights into internal and external factors influencing service quality and enables the formulation of targeted solutions.

Overall, this study is expected to provide practical contributions to the banking sector in improving service quality, strengthening brand awareness, and enhancing user experience based on empirical data (Kaur, 2021). Additionally, it contributes academically to the development of studies on the application of sentiment analysis in digital banking, particularly in the context of strategic decision-making oriented toward customer needs and perceptions.

METHOD

Descriptive Method

The descriptive method is used to scientifically describe actual phenomena, including conditions, inter-component relationships, and prevailing opinions. This method aims to provide an objective understanding of the problem under investigation.

Exploratory Method

The exploratory method is employed to understand patterns, characteristics, and data structures before conducting in-depth analysis. In this study, exploration includes word frequency analysis, emoji usage, and sentiment distribution to guide preprocessing strategies and the selection of relevant features.

Quantitative Research Method

The quantitative method focuses on the objective collection and analysis of numerical data to identify relationships between variables and test hypotheses (Mitchell, 1997; Monika et al., 2022). This study utilizes both primary and secondary data, allowing the results to be measured, statistically validated, and generalized more broadly.

RESULT AND DISCUSSION

Data collection was conducted through crawling processes using IGCommentExporter, InsCommentExport, and TTCommentExporter modules to extract comments from Instagram and TikTok in CSV or Excel format (Sharma & Kaushik, 2023). Extraction was performed based on the URL of each official Bank post during the period from August 31, 2024, to August 31, 2025, with module selection adjusted according to the number of comments on each post. A total of 10,198 comments were collected, comprising 4,217 from Instagram and 5,981 from TikTok.

	A	B	C	D	E
	User ID	Username	Comment ID	Comment	Comment Tim
1	38317571014	zzzscraaft	="17937585944855324	@wtherery iyaa nunggu gajian lumayan banget nih promonya jujur!	7/21/2024, 1:15:42
2	49770779028	zzungokong2	="17987169425525163	@bimskyy07 lemes bgt jempolnya	7/21/2024, 2:20:47
3	19812411723	zyertxx	="18282313426173959	min gaada rencana mau nambahin barcode payment pake qris kah? soalnya kalau setiap ke indomare	11/13/2024, 7:39:5
4	19812411723	zyertxx	="18049437032006413	@blubybcadigital kira kira kapan ya perilisan nya? bulan ini kah?	11/14/2024, 1:25:5
5	9289176634	zyanmansyah	="18353421274136690	khusus tarik tunai di atm biaya adminnya gratis dong kak td pagi mau narik tunai pas, kurang ada biaya	9/25/2024, 10:56:0
6	31841405659	zilong6811	="18050097439972864	Aman gak sih nabung di blu ini ?	11/24/2024, 10:26:
7	35930747	zeliladya	="17914075730914253	min gabisa tf mulu nii dari pagi , lagi gangguan apa gimana?	10/9/2024, 4:32:04
8	3500705701	zaodee	="18035243588020822	Blu tolong dong aplikasinya jangan kebanyakan animasi, loading, jadi berat dan bertele-tele bgt klo lgi	7/13/2024, 8:50:32
9	10197036288	zan.elsigli	="17862150540209323	🖤	8/22/2024, 10:56:2
10	51259877524	zaliwaaa	="18047076337681519	min adain promo bareng cinepolis lg dong, di dom ku blm ada cgv😔😔😔	7/11/2024, 12:32:0
11	7293520412	zakighufron9	="18071882398559384	Pernikahan bisa dikatakan ibadah terpanjang yang mesti dijalani oleh pasangan yang sudah sah dalan	8/9/2024, 2:05:33
12	7293520412	zakighufron9	="17990438369676438	Jawaban No. 1. C. Zumba x Dance Party 2. C. bluGether must win because i want to prepare to join wo	8/13/2024, 6:32:32
13	7293520412	zakighufron9	="18043911286767045	Keresahan seputar karir seiring berjalannya waktu dan bertambahnya usia memang bukan suatu hal y	8/6/2024, 9:49:22
14	7293520412	zakighufron9	="18358191145105619	Teman adalah bagian dari partner yang datang, hilang terseleksi oleh situasi, terkadang muncul lagi. S	8/7/2024, 5:44:07

Figure 1. The collected social media user comments

Social media text presents unique challenges for information retrieval and sentiment classification due to its informal language, high variability, and contextual dependency, requiring appropriate preprocessing and representation techniques (Pak & Paroubek, 2010; Vechtomova, 2009).

Data Preprocessing Phase

The initial stage of data preprocessing (data preparation) aims to produce clean and structured data to support further processing. The following steps were conducted during data preparation:

1. Data Import

The first step in data processing involved reading the raw data stored in Excel format and loading it into a structured data frame using Python in Jupyter Notebook. The file rawdata.xlsx was successfully imported, containing a total of 10,198 rows and 4 columns.

```
df = pd.read_excel("rawdata.xlsx")
df.head()
```

	username	komentar	tanggal
0	bibiwniu	kalo join acaranya bisa join blu juga ngga min	20-8-2024 7:43:33 PM
1	farelgilangg	tarik tunai di atm bca kena admin,kuota gratis...	30-11-2024 1:58:09 PM
2	andihdyat	Done	30-8-2024 8:55:58 PM
3	gianti.meisitha	Sama.. saya barusan tf ke bca.. bacaannya berh...	02-12-2024 1:01:46 PM
4	gianti.meisitha	Terus dananya nyangkut dimana saldo blu udh k...	02-12-2024 1:08:42 PM

```
Jumlah record: 10198
Ukuran dataset (baris, kolom): (10198, 4)
```

Figure 2. The imported data within Jupyter Notebook

2. Handling Missing Values

Missing value processing is conducted to ensure the dataset meets standardized quality, making the data complete, consistent, and reliable for analysis. Incomplete data can introduce bias and reduce analytical accuracy; therefore, missing entries are removed or otherwise addressed to maintain dataset integrity.

```
Jumlah data sebelum menghapus missing value: 10198
Jumlah data setelah menghapus missing value: 10194
```

[6]:

	komentar	UserID	username	datetime
0	@acacantiktpiboonq lopppp ❤️	1554192	ayudiac	03 April 2025 18:27:43
1	@ajrunrka kerennnn	1554192	ayudiac	03 April 2025 18:27:54
2	@putradjohan_aloke wakakakaka	1554192	ayudiac	03 April 2025 18:28:12
3	@arbisono_ tosss	1554192	ayudiac	03 April 2025 19:49:27
4	@poppyrosaline ih asik amattt	1554192	ayudiac	03 April 2025 19:49:34

Figure 3. Processing of Empty Data Removal

3. Removing Duplicates

Removing duplicates eliminates identical entries to prevent repetition of comments that may have been collected multiple times during initial data extraction.

```
Jumlah data awal: 10194
Setelah menghapus missing value: 10194
Setelah menghapus duplikasi: 9708
```

]:

	komentar	UserID	username	datetime
0	@acacantiktpiboonq lopppp ❤️	1554192	ayudiac	03 April 2025 18:27:43
1	@ajrunrka kerennnn	1554192	ayudiac	03 April 2025 18:27:54
2	@putradjohan_aloke wakakakaka	1554192	ayudiac	03 April 2025 18:28:12
3	@arbisono_ tosss	1554192	ayudiac	03 April 2025 19:49:27
4	@poppyrosaline ih asik amattt	1554192	ayudiac	03 April 2025 19:49:34
5	Yeaaay Done	3566320	resty_wulan	31 August 2024 23:42:11

Figure 4. Process of Duplicate Removal

4. Case Folding

At this stage, text is standardized by converting all characters to lowercase, ensuring that every word or sentence in the text attribute follows a consistent format to support accurate data analysis.

	komentar
0	@acacantiktpiboonq lopppp ❤️
1	@ajrunrka kerennnn
2	@putradjohan_aloke wakakakaka
3	@arbisono_ tosss
4	@poppyrosaline ih asik amattt
5	yeaaay done
6	slalu top up blu dari bca, tujuannya buat nabu...
7	@mevlbourne tmrw juga gratis, unlimited lagi ????
8	best thing i've ever invested in this year 💙
9	@blubybcadigital sebagai masukan tolong kasih ...

Figure 5. Letter Standardization Process

5. Lemmatization

Lemmatization is the process of normalizing text by reducing words to their base or lemma form, allowing variations of words with similar meanings to be unified in the analysis. Applying lemmatization enhances accuracy in natural language processing (NLP) by representing each word in its most grammatically representative form.

[13]:

	komentar	komentar_stem
0	@acacantiktpiboonq lopppp ❤️	acacantiktpiboonq lopppp
1	@ajrunrka kerennnn	ajrunrka kerennnn
2	@putradjohan_aloke wakakakaka	putradjohan aloke wakakakaka
3	@arbisono_ tosss	arbisono tosss
4	@poppyrosaline ih asik amattt	poppyrosaline ih asik amattt
5	yeaaay done	yeaaay done
6	slalu top up blu dari bca, tujuannya buat nabu...	slalu top up blu dari bca tuju buat nabung di ...
7	@mevlbourne tmrw juga gratis, unlimited lagi ????	mevlbourne tmrw juga gratis unlimited lagi

Figure 6. Text Normalization Process (Lemmatization)

6. Tokenization

Tokenization involves splitting text into smaller units, such as words, phrases, or symbols. This process transforms lengthy text into a form that is easier to process and analyze. Tokens, particularly adjectives, verbs, or domain-specific terms, are then available for further sentiment analysis.

	komentar_stem	tokens
0	acacantiktpiboonq lopppp	[acacantiktpiboonq, lopppp]
1	ajrunrka kerennnn	[ajrunrka, kerennnn]
2	putradjohan aloke wakakakaka	[putradjohan, aloke, wakakakaka]
3	arbisono tosss	[arbisono, tosss]
4	poppyrosaline ih asik amattt	[poppyrosaline, ih, asik, amattt]
5	yeaaay done	[yeaaay, done]

Figure 7. Tokenization Process

7. Stopwords Removal

Stopwords removal eliminates common words that do not contribute significantly to the analytical context, improving efficiency and accuracy in text processing, especially for sentiment analysis.

	tokens	tokens_clean
0	[acacantiktpiboonq, lopppp]	[acacantiktpiboonq, lopppp]
1	[ajrunrka, kerennnn]	[ajrunrka, kerennnn]
2	[putradjohan, aloke, wakakakaka]	[putradjohan, aloke, wakakakaka]
3	[arbisono, tosss]	[arbisono, tosss]
4	[poppyrosaline, ih, asik, amattt]	[poppyrosaline, ih, asik, amattt]
5	[yeaaay, done]	[yeaaay, done]
6	[sialu, top, up, blu, dari, bca, tuju, buat, n...	[sialu, top, up, blu, bca, tuju, buat, nabung,...
7	[mevlbourne, tmrw, juga, gratis, unlimited, lg]	[mevlbourne, tmrw, gratis, unlimited, lg]
8	[best, thing, i, ve, ever, invested, in, this,...	[best, thing, i, ve, ever, invested, in, this,...
9	[blubybcadigital, bagai, masuk, tolong, kasih,...	[blubybcadigital, bagai, masuk, kasih, kuota, ...
10	[ide, bisa, pakai, untuk, hadap, tipu]	[ide, pakai, hadap, tipu]
11	[app, down, lagi, down, lagi, so, inconvenient]	[app, down, down, so, inconvenient]

Figure 8. Stopword Removal Process

8. Stemming

Stemming reduces words to their base form. For example, “membaca,” “pembaca,” and “dibaca” are all converted to “baca.” This step minimizes unnecessary word variations in the text corpus, allowing analysis to focus on core meanings consistently.

	komentar	tokens_stemmed
0	@acacantiktpiboonq lopppp 💗	[acacantiktpiboonq, lopppp]
1	@ajrunrka kerennnn	[ajrunrka, kerennnn]
2	@putradjohan_aloke wakakakaka	[putradjohan aloke, wakakakaka]
3	@arbisono_ tosss	[arbisono, tosss]
4	@poppyrosaline ih asik amattt	[poppyrosaline, ih, asik, amattt]

Figure 9. Word Base-Form Normalization Process

9. Word Translation

In sentiment analysis, word translation converts Indonesian text into English to enable processing with models trained on English-language datasets. This step improves sentiment classification accuracy by leveraging more comprehensive English-language models and resources.

	komentar	tweet_english
0	kalo join acara join blu ngga min	if you join the event join blu no min
1	tarik tunai atm bca kena adminkuota gratis nya...	cash withdrawal atm bca get free admission its...
2	done	English language
3	sama barusan tf bca baca hasil saldo udh poton...	same as just now tf bca read the results of th...
4	terus dana nyangkut saldo blu udh potong tp ki...	then the fund relates to the balance of blu ud...

Figure 10. Text Data Translation Stages

These preprocessing steps are standard practices in natural language processing to improve data consistency, reduce noise, and enhance the accuracy of sentiment classification (Loria, 2018; Martin & Jurafsky, 2009).

10. Splitting Training and Testing Datasets

The dataset is divided into 80% training and 20% testing data, using stratified parameters to maintain balanced label distribution. This ensures the model learns patterns accurately and can be evaluated for generalization performance, providing valid and reliable results on new data.

```
# ================================
# PEMBAGIAN DATA TRAINING & TESTING
# ================================

from sklearn.model_selection import train_test_split
import pandas as pd

# Ambil fitur (X) dan label (y)
X = data['tweet_clean']          # kolom teks yang sudah dibersihkan
y = data['klasifikasi']          # label sentimen (Positif, Negatif, Netral)

# Bagi data menjadi Training dan Testing
X_train, X_test, y_train, y_test = train_test_split(
    X, y,
    test_size=0.2,
    stratify=y,
    random_state=42
)
```

Figure 11. Training and Testing Data Configuration

```
Jumlah Data Total   : 8942
Jumlah Data Training: 7153
Jumlah Data Testing : 1789

Distribusi Label Training:
klasifikasi
Netral     0.603523
Positif    0.264225
Negatif    0.132252
Name: proportion, dtype: float64

Distribusi Label Testing:
klasifikasi
Netral     0.603130
Positif    0.264394
Negatif    0.132476
Name: proportion, dtype: float64
```

Figure 12. Distribution of Training and Testing Data Based on Sentiment Labels

11. Data Vectorization

Data vectorization is performed using Bag of Words (BoW) and TF-IDF methods to convert preprocessed, multilingual, and normalized text into numerical matrix representations (Bishop & Nasrabadi, 2006; Hastie et al., 2005). TF-IDF extracted 5,000 features with the 30 most significant words, such as want, card, and just, which carry high weights even if not present in all comments.

```
Jumlah fitur TF-IDF: 5000

=== Matrix Top 30 TF-IDF ===
        want        card        just         ok     account     error  \
0  158.060942  127.581448  109.472653  109.22808  108.563247  98.06392

       admin         dm      check      pay  ...      using        blu  \
0  97.363971  93.347458  88.044168  86.0041  ...  62.488801  62.064864

      debit  transaction       free         oh       long       used  \
0  59.85721    59.109404  58.815837  58.291536  57.438016  57.359711

    balance      money
0  56.72286  56.568355

[1 rows x 30 columns]
```

Figure 13. Text Representation Using TF-IDF

BoW highlighted the most frequently occurring words, including want, card, just, and account, reflecting topics largely related to digital financial services, transactions, and banking account issues.

```
Jumlah fitur BoW: 5000

=== Matrix Top 30 Bag of Words (tanpa stopwords) ===
   want  card  just  account   im  ive  dont  pay  like  using  ...  long  \
0   808   594   550      506  361  359   357  356   322    285  ...   226

   admin  balance  debit  check  used  open  number   ok  transaction
0    224      223    223    214   212   203     199  198          194

[1 rows x 30 columns]
```

Figure 14. Text Representation Using the Bag-of-Words (BoW) Model

Sentiment Classification Data Processing

After preprocessing, the dataset was reduced from 10,198 to 8,972 comments, as some data were identified as duplicates, non-informative, or consisting solely of symbols and non-alphabetic characters. The remaining dataset is considered more valid and standardized, making it optimal for sentiment classification to produce accurate and reliable analyses.

1. Naïve Bayes Classifier
   a. Model Accuracy Evaluation

Sentiment classification using the Naïve Bayes method was performed on the preprocessed dataset, including model training and testing to recognize sentiment patterns in text. The model achieved an accuracy of 84.29%, which falls within the typical performance range of Naïve Bayes for sentiment analysis (Fitri et al., 2019). The classification report indicated that the model was most accurate in detecting neutral sentiment, while performance for negative sentiment remained limited due to low recall, and positive sentiment exhibited relatively good performance. Weighted averages for precision, recall, and F1-score were 0.85, 0.84, and 0.83, respectively, reflecting balanced overall performance. The confusion matrix revealed misclassifications, particularly for the

negative class, which was often classified as neutral or positive, suggesting contextual similarities between sentiment categories.

From a theoretical perspective, Naïve Bayes is known to rely on probabilistic assumptions that may limit its ability to capture complex contextual dependencies in text data, which can affect classification performance in sentiment analysis tasks (Kang et al., 2012; Murphy, 2012).

```
Akurasi Test: 0.8429290106204583

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.86      0.39      0.53       237
     Netral        0.88      0.94      0.91      1079
     Positif       0.77      0.86      0.81       473

    accuracy                           0.84      1789
   macro avg       0.83      0.73      0.75      1789
weighted avg       0.85      0.84      0.83      1789
```

Figure 15. Model Performance Evaluation Using the Classification Report

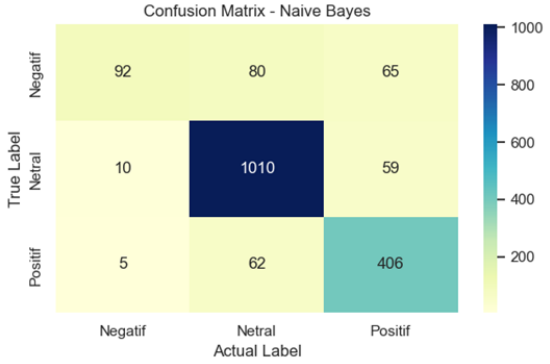

Figure 16. Naive Bayes Model Evaluation Using the Confusion Matrix

b. Naïve Bayes Sentiment Processing

Naïve Bayes classification results showed that most user comments were neutral (5,422), followed by positive sentiment (2,363) and negative sentiment (1,183). This distribution indicates that user perceptions of ABC's services tend to be stable, with positive sentiment exceeding negative, reflecting a relatively satisfactory user experience.

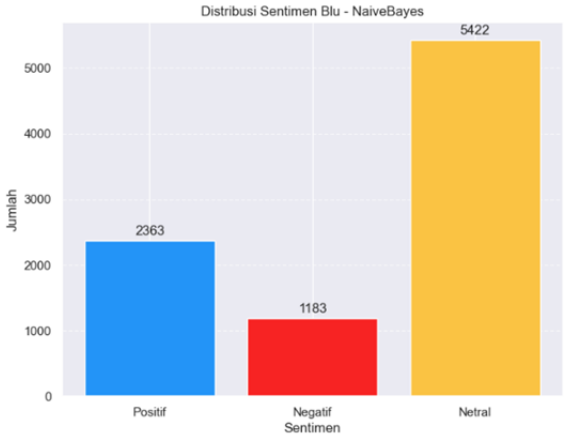

Figure 17. Sentiment Distribution Chart Based on Naïve Bayes Classification

WordCloud visualizations were used to illustrate the most influential words in shaping sentiment based on frequency and probabilistic weight. For positive sentiment, terms such as free, first, account, ok, safe, and good dominated, reflecting perceptions of convenience, security, and benefits of the digital banking service. Negative sentiment was dominated by words such as long, cant, time, transaction, and fee, highlighting complaints related to service duration, transaction issues, and technical disruptions, including application and international service problems. This WordCloud representation based on Naïve Bayes is considered more accurate due to its probabilistic weighting, which reflects user sentiment tendencies more precisely.

Figure 18. Positive Sentiment Word Cloud for Naïve Bayes

Figure 19. Naïve Bayes Negative Sentiment Word Cloud

2. Support Vector Machine (SVM) Classification
   a. Model Accuracy Evaluation

The Support Vector Machine (SVM) model demonstrated excellent performance in sentiment classification, achieving an overall accuracy of 92.11%. According to the classification report, neutral sentiment exhibited the best performance (precision 0.92; recall 0.99), followed by positive sentiment (precision 0.94; recall 0.87), while negative sentiment showed good precision but higher misclassification rates (precision 0.90; recall 0.70). Macro and weighted average scores were approximately 0.92, indicating consistent performance across all classes.

```
Akurasi Test: 0.9211850195640022

Classification Report:
              precision    recall  f1-score   support

     Negatif       0.90      0.70      0.79       237
      Netral       0.92      0.99      0.95      1079
     Positif       0.94      0.87      0.91       473

    accuracy                           0.92      1789
   macro avg       0.92      0.85      0.88      1789
weighted avg       0.92      0.92      0.92      1789
```

Figure 20. SVM Model Classification Report

The confusion matrix further confirmed these results, showing that most data in each category were correctly classified. However, the negative class had a higher misclassification rate, with only 166 out of 237 negative comments correctly classified, compared to 1,070 out of 1,079 neutral comments and 412 out of 473 positive comments.

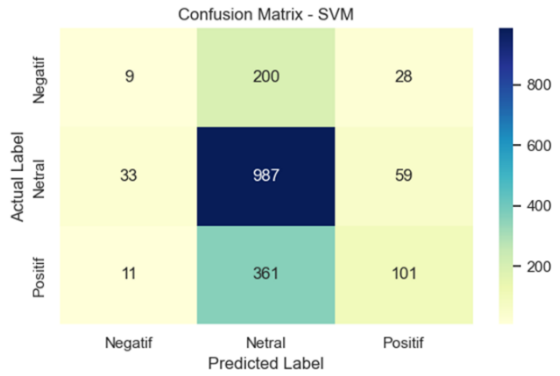

Figure 21. SVM Model Evaluation Using the Confusion Matrix

b. SVM Sentiment Processing

SVM classification results indicated that neutral sentiment dominated with 5,671 comments, followed by positive sentiment (2,229) and negative sentiment (1,042). This finding suggests that the majority of responses are informative or neutral, with positive reviews exceeding negative, reflecting generally favorable public perception of the ABC application.

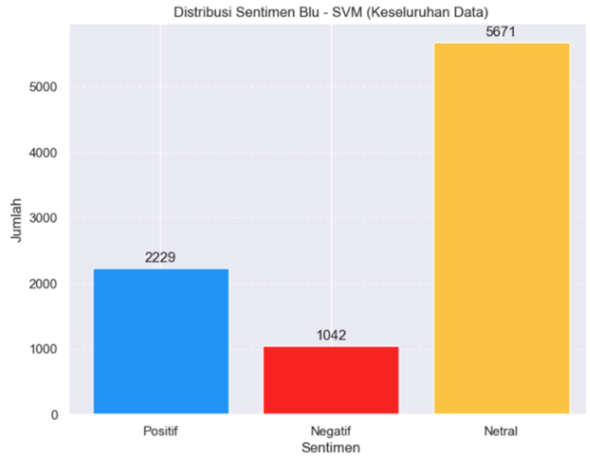

Figure 22. Sentiment Distribution Chart Based on SVM Classification

c. SVM Word Visualization

WordCloud visualizations from the SVM classification highlight dominant terms reflecting user perceptions of digital banking services. Positive sentiment was dominated by words such as free, card, first, safe, better, and good, indicating appreciation for transaction convenience, security, and overall app experience. Frequently occurring terms like transfer, account, pay, and saving emphasize that transaction features are the most valued aspects. Conversely, negative sentiment was dominated by words such as time, long, cant, transaction, bank, account, and slow, indicating user complaints primarily related to processing time, transaction difficulties, account access limitations, and suboptimal application performance.

Figure 23. Positive Sentiment Word Cloud of SVM

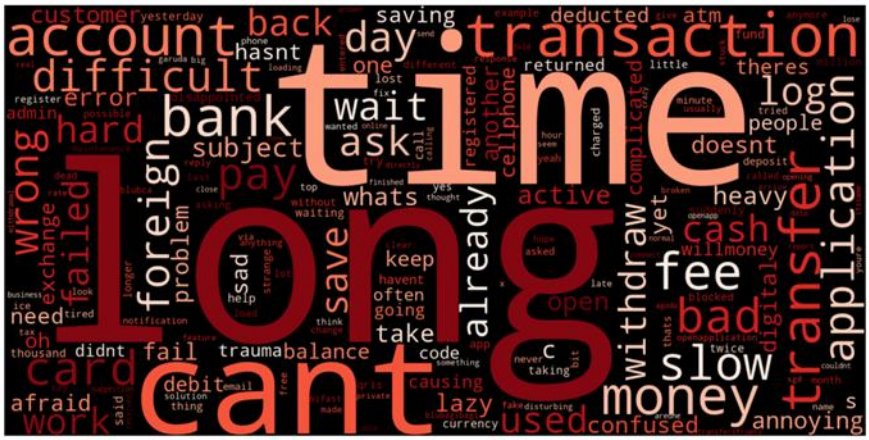

Figure 24. Negative Sentiment Word Cloud of SVM

Gap Identification between Naïve Bayes and SVM Classification

Analysis of 8,942 comments revealed that 442 comments (4.94%) were classified differently between Naïve Bayes and SVM, reflecting variations in feature interpretation due to Naïve Bayes' probabilistic approach and SVM's hyperplane separation.

Table 1. Perbandingan Hasil Analisis Sentimen Naïve Bayes dan SVM

		Support Vector Machine (SVM)			Grand Total
		Negatif	Netral	Positif	
Naïve Bayes	Negatif		146	55	201
	Netral	7		23	30
	Positif	52	159		211

TOTAL	**59**	**305**	**78**	**442**

Overall performance evaluation indicates that SVM outperformed Naïve Bayes, achieving an accuracy of 0.9234, precision 0.92, recall 0.86, and F1-score 0.88, compared to Naïve Bayes with accuracy 0.8429, precision 0.83, recall 0.73, and F1-score 0.75 (Asriana et al., 2024; Jannah & Kusnawi, 2024; Korovkinas et al., 2017). The most notable performance difference was observed in negative sentiment, where SVM achieved a recall of 0.70 versus 0.39 for Naïve Bayes, demonstrating SVM's superior capability in detecting user complaints and reducing false negatives.

Table 2. Comparison of the Performance of Naïve Bayes and SVM Models

Model Performance Comparison: SVM and Naïve Bayes				
Method	**Accuracy**	**Precision (Macro Avg)**	**Recall (Macro Avg)**	**F1-Score (Macro Avg)**
Naïve Bayes	0,8429	0,83	0,73	0,75
SVM	0,9234	0,92	0,86	0,88

Per-Class Sentiment Comparison (SVM and Naïve Bayes)					
Model	**Sentiment**	**Precision**	**Recall**	**F1-Score**	**Support**
Naïve Bayes	Negative	0,86	0,39	0,53	237
	Neutral	0,88	0,94	0,91	1.079
	Positive	0,77	0,86	0,81	473
SVM	Negative	0,91	0,7	0,79	237
	Neutral	0,92	0,99	0,95	1.079
	Positive	0,94	0,88	0,91	473

Consistency in handling imbalanced data. Naïve Bayes' lower performance is influenced by its feature independence assumption, which does not align well with the contextual and diverse nature of social media text. Consequently, SVM is recommended as the final model due to its higher accuracy and stability, particularly in detecting actionable user criticism.

Strategic Analysis Using SWOT

The SWOT analysis identifies user sentiment toward digital banking services through four dimensions: Strengths, Weaknesses, Opportunities, and Threats (David et al., 2017). This mapping provides insight into internal and external factors affecting user perception and experience, serving as a strategic basis for enhancing service quality and competitiveness (Utami & Hidayat, 2024; Utomo & Hidayat, 2023).

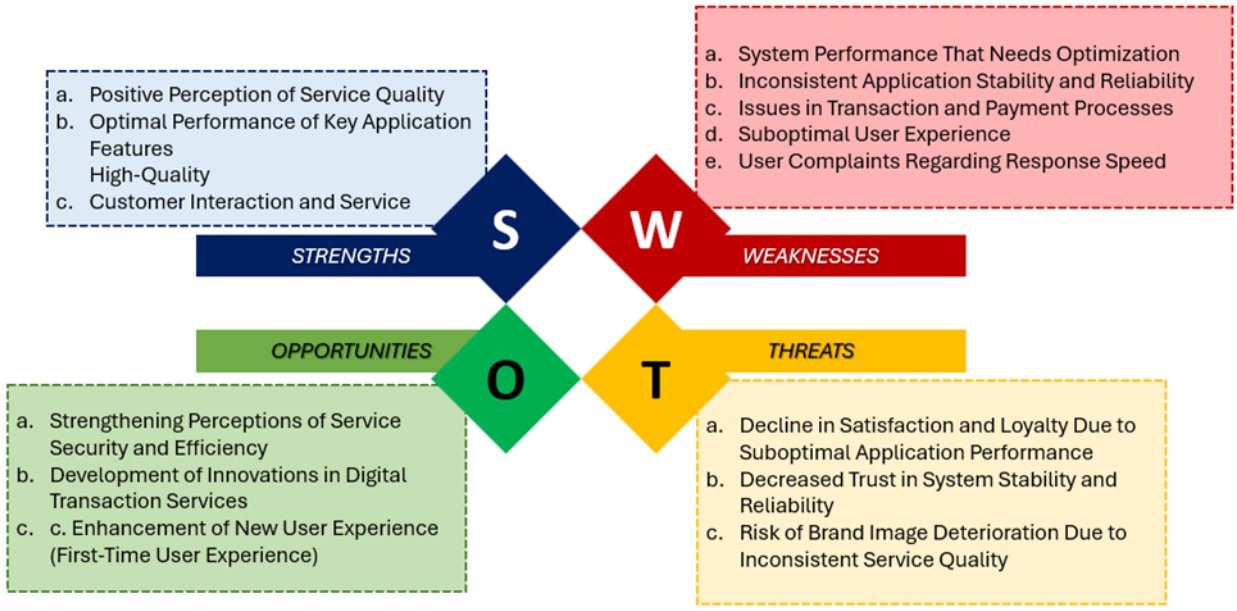

Figure 25. SWOT Analysis of Digital Banking Services

- Strengths: Positive perceptions of service quality are reflected in terms such as free, ok, good, and safe, highlighting convenience, efficiency, and transaction security. Core feature performance is demonstrated by words like card, transfer, and account, indicating functional and valuable application features. Quality customer interaction is represented by admin and message, reflecting responsive and communicative service experiences.
- Weaknesses: System performance needs optimization, indicated by long, slow, and time. Application stability and reliability issues are revealed through cant, log, wrong, and difficult, reflecting technical challenges with login and transactions. Transaction and payment problems are suggested by transfer, transaction, and pay, while poor user experience is indicated by bad and application. Slow customer support response is reflected by ask and already.
- Opportunities: Strengthening perceptions of security and efficiency (safe, save), developing faster, personalized, and integrated digital transaction features (transfer, card), and enhancing onboarding experiences for new users (good, first) can improve retention and satisfaction.
- Threats: Declining satisfaction and loyalty are indicated by long, slow, and wait. System reliability concerns (error, log) may cause user migration, while negative perceptions (application, transaction) can undermine brand reputation and competitiveness.

Root Cause Analysis

1. Fishbone Diagram

Fishbone analysis, derived from identified weaknesses in SWOT and reinforced by elements of Strengths, Opportunities, and Threats, highlights suboptimal application performance, including slow transactions, system instability, login issues, and unsatisfactory user interactions. Key causes are primarily linked to technological and infrastructure limitations, inefficient operational processes, and insufficient customer support, while interaction design and management policies exacerbate user dissatisfaction. Main causal categories include People, Process, Technology, System Integration, and Policy Management, enabling a comprehensive examination of service issues from multiple perspectives.

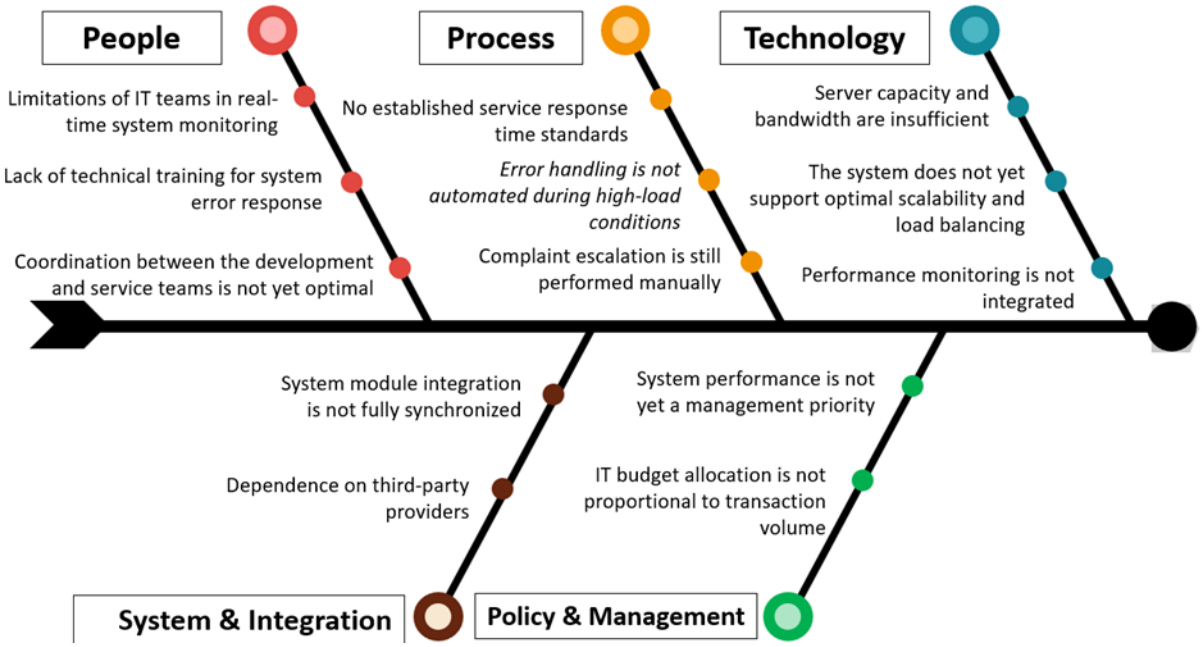

Figure 26. Fishbone Diagram of Suboptimal System Performance and Service Speed

The primary issue concerns system performance and service speed, which do not meet optimal standards. This stems from Weakness points (a) and (e) and Threat point (a), reflecting insufficient system performance and slow service response. Root causes include technological infrastructure constraints, suboptimal system architecture, and the absence of standardized response times. Additionally, limited real-time performance monitoring and delayed incident handling by human resources contribute to suboptimal service delivery.

2. Whys Analysis

The 5 Whys Analysis approach was conducted to systematically identify and trace the root causes of non-compliance in handling technical incidents. This section presents a scenario-based 5 Whys analysis derived from Fishbone 1: Suboptimal System Performance and Service Speed.

a. Suboptimal System Performance and Service Speed

Scenario 1: Slow Transaction Response Time

The 5 Whys analysis for this scenario was carried out to identify the causes of delayed response times in digital banking transactions. The root causes were found to be server capacity limitations and suboptimal implementation of load balancing systems, which hinder the system's ability to process transactions efficiently and promptly.

Table 3. Five Whys Analysis: Slow Transaction Response Time

Main Problem	
The digital banking application exhibits slow response times when processing financial transactions.	
First Why	Because the server capacity is not yet able to handle a surge in simultaneous users.
Second Why	Because the load balancing system has not been optimized to distribute workloads evenly.
Third Why	Because the system architecture does not yet support dynamic, on-demand scalability (auto-scaling).

Fourth Why	Because there is no policy for managing infrastructure based on cloud resource management.
Fifth Why	Because prioritizing system performance improvements has not yet become a strategic agenda for IT management.
Solusi	Implementing an adaptive auto-scaling and load balancing system, supported by strategic policies to strengthen cloud-based infrastructure, to ensure service performance and speed.

Follow-up Plan	
Action	**Coordinator**
Evaluation of server capacity and application usage patterns	IT Ops & Infrastructure
Implementation of adaptive load balancing	IT Ops & Infrastructure
Integration of auto-scaling policies	IT Ops & Infrastructure

Formulation of Improvement Recommendations

The formulation of comprehensive improvement recommendations, based on in-depth analysis and aligned with user-centered design principles, is necessary to ensure the creation of a more consistent, efficient, and sustainable service design. Through a systematic approach, these recommendations aim to enhance the quality of digital banking products and improve the overall user experience.

1. Suboptimal System Performance and Service Speed

Scenario 1: Slow Transaction Response Time

Slow transaction response times in digital banking services indicate limitations in system capacity to handle surges in user demand. This condition has the potential to reduce user experience quality and negatively affect perceptions of service reliability. The recommended improvements focus on enhancing system performance through infrastructure capacity optimization, implementation of adaptive scalability, and improved workload distribution, thereby enabling digital transactions to operate more quickly, stably, and responsively under various usage scenarios.

Table 4. Slow Transaction Response Time Recommendations

Follow-up Plan			
Action	**Owner**	**Priority**	**Status**
Evaluation of server capacity and application usage patterns	IT Ops & Infrastructure	High	Open
"Evaluating server capacity and application usage patterns is a strategic effort to ensure the system can handle workloads optimally. This high-priority activity aims to identify potential bottlenecks, determine capacity enhancement needs, and serve as the basis for implementing adaptive load balancing, thereby maintaining the stability and performance of digital banking services."			
Management Response Management approves the implementation of server capacity and application usage pattern evaluations as a critical step in maintaining service performance and stability. Management emphasizes that the results of this evaluation must be analyzed comprehensively to ensure that capacity enhancement recommendations, including the implementation of adaptive load balancing, can be executed effectively and aligned with operational needs as well as projected growth of digital banking services. (Estimated Resolution Date: January 1, 2026)			

Implementation of adaptive load balancing	IT Ops & Infrastructure	High	Closed
"The implementation of adaptive load balancing enables dynamic distribution of system workloads according to traffic conditions and resource capacity, thereby maintaining service performance at an optimal level. This approach also enhances system resilience by minimizing the risk of bottlenecks and ensuring service continuity across various usage scenarios." **Management Response** Management has implemented adaptive load balancing as a strategic measure to enhance service stability and performance. Management has also applied this mechanism alongside regular monitoring and measurable technical evaluations, ensuring that the effectiveness of workload distribution can be continuously improved and aligned with the operational needs of digital banking services.			
Implementation of server auto-scaling policy integration	IT Ops & Infrastructure	Medium	Open
The integration of auto-scaling policies is carried out to enhance the flexibility and availability of IT services by automatically adjusting server capacity according to workload fluctuations. This initiative by the IT Infrastructure and Operations team aims to maintain service performance, reduce the risk of downtime, and improve the efficiency of resource and cost utilization. **Management Response** Management approves the implementation of auto-scaling policies as a measure to achieve adaptive and efficient capacity enhancement. Management also emphasizes the importance of establishing clear configuration parameters and conducting regular evaluations to ensure that the capacity adjustment mechanism operates optimally and supports the stable continuity of digital banking services.			

Evaluation and Validation of Recommendations

Based on the recommendations discussed with management, the majority have been implemented (status: CLOSE) to enhance the quality of digital banking services. Some recommendations, still marked as OPEN, such as improving human resource capacity and infrastructure monitoring, indicate that implementation processes are ongoing, with management demonstrating a commitment to executing them progressively.

Table 5. Management Response Regarding Slow Transaction Response Time

No	Results of the Evaluation and Validation of Follow-Up Implementation
1	**Server Capacity and Application Usage Evaluation** Evaluating server capacity and application usage patterns is a strategic effort to ensure the system can handle workloads optimally. This high-priority activity aims to identify potential bottlenecks, determine capacity enhancement needs, and serve as the basis for implementing adaptive load balancing, thereby maintaining the stability and performance of digital banking services. **Management Response (IT Ops & Infrastructure)** Management approved the implementation of server capacity and application usage evaluations as a critical step to maintain service performance and stability. Management emphasized that the evaluation results must be analyzed comprehensively to ensure capacity enhancement recommendations, such as adaptive load balancing, can be implemented effectively and aligned with operational requirements and projected growth of digital banking services (Estimated Completion: January 1, 2026).

	Recommendation Evaluation The recommendation to evaluate server capacity and analyze application usage patterns is considered relevant to ensure service stability and performance. Observations indicate that this recommendation is valid and implementable, as it supports the identification of potential bottlenecks, determination of capacity enhancement needs, and the measured implementation of adaptive load balancing. Management has approved the execution, emphasizing the need for comprehensive evaluation of the results to implement capacity improvements effectively and in alignment with projected service growth.
2	**Integration of Auto-Scaling Policy** The integration of an auto-scaling policy aims to enhance IT service flexibility and availability by automatically adjusting server capacity in accordance with workload fluctuations. This initiative from the IT Infrastructure and Operations team seeks to maintain service performance, reduce downtime risk, and improve resource and cost efficiency. **Management Response (IT Ops & Infrastructure)** Management approved the implementation of the auto-scaling policy as an adaptive and efficient capacity enhancement measure. Management also stressed the importance of defining clear configuration parameters and conducting regular evaluations to ensure the capacity adjustment mechanism operates optimally and supports the continuity of digital banking services. **Recommendation Evaluation** The auto-scaling policy recommendation is deemed appropriate and strategically relevant for increasing flexibility, availability, and efficiency in handling workload fluctuations. Automatic server capacity adjustments are considered valid and implementable, as they help maintain service performance, minimize downtime risk, and optimize resource utilization. Management approved this recommendation, emphasizing the necessity of clear configuration parameters and periodic evaluations to ensure the auto-scaling mechanism functions effectively and consistently supports digital banking service continuity.

CONCLUSION

The evaluation results reveal that the Support Vector Machine (SVM) model outperforms the Naïve Bayes algorithm in sentiment classification, demonstrating the superior ability of SVM to optimally separate data classes within the scope of this analysis. Regarding digital banking services, sentiment analysis indicates that the majority of users hold neutral a nd positive perceptions, reflecting a relatively stable level of user satisfaction. Both models, Naïve Bayes and SVM, highlight a higher proportion of positive comments compared to negative ones, with a significant amount of neutral sentiment observed in the dataset. The data processing methodology, which incorporates steps such as text standardization, word normalization, and language translation, contributes significantly to improving sentiment analysis accuracy and ensures the relevance and validity of the data used. This approach is essential for reliable model performance and the generation of robust analytical results. Despite this, the performance of digital banking systems still exhibits certain areas that require improvement, particularly related to slow transaction processing, system instability, and login issues, all of which affect the overall user experience. In addition to technological and infrastructural factors, inefficient operational processes also play a substantial role in these challenges.

To improve sentiment analysis accuracy and comprehensiveness, it is recommended to conduct regular data collection from various social media platforms and to perform systematic data cleaning (pre-processing) to obtain more accurate and representative results. Moreover, the development of advanced sentiment classification models,

particularly those utilizing deep learning algorithms, can help reduce misclassification rates and offer a more precise representation of user perceptions. Continuous model training with updated data is also essential. Additionally, investing in training and enhancing the capabilities of human resources, especially in data analysis, communication, and IT utilization, is crucial to strengthening internal insights and improving service quality. This competency development will foster evidence-based and data-driven decision-making. Finally, enhancing system specifications and technological infrastructure is vital to ensure faster and more reliable digital banking transactions, which would improve the overall user experience. Prioritizing, evaluating, and periodically implementing infrastructure improvements should be a key focus for continued progress.

REFERENCES

Alsemaree, O., Alam, A. S., Gill, S. S., & Uhlig, S. (2024). Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions. *Heliyon*, *10*(9).

Asriana, A., Pratiwi, O. N., & Fakhrurroja, H. (2024). Analisis Sentimen Pada Media Sosial Twitter Menggunakan Support Vector Machine (SVM) Terhadap Kata Kunci# TheLinkinJKT. *EProceedings of Engineering*, *11*(4).

Axhiu, M., Veljanoska, F., Ciglovska, B., & Husejni, M. (2014). The Usage of Sentiment Analysis for Hearing the Voice of the Customer and Improving Businesses. *Journal of Educational and Social Research*, *4*. https://doi.org/10.5901/jesr.2014.v4n4p401

Bhardwaj, A., Bharany, S., & Kim, S. (2024). Fake social media news and distorted campaign detection framework using sentiment analysis & machine learning. *Heliyon*, *10*(16).

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, Issue 4). Springer.

Chen, Y., Wang, Z., Yuan, G., & Huang, L. (2017). An overview of online based platforms for sharing and analyzing electrophysiology data from big data perspective. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *7*(4), e1206.

David, F. R., David, F. R., & David, M. E. (2017). *Strategic management: concepts and cases: A competitive advantage approach* (Vol. 127). Pearson New Jersey.

Divya, B. (2019). *Performance Management and its Counseling in Manufacturing Units*.

Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, *161*, 707–714.

Fauzi, A., Kom, S., Kom, M., Yunial, A. H., Kom, S., & Kom, M. (2024). *Analisis Sentimen (Sentiment Analysis): Evaluasi Sentimen Layanan Dataset Twitter US Airline*. CV. Bintang Semesta Media.

Fitri, V. A., Andreswari, R., & Hasibuan, M. A. (2019). Sentiment analysis of social media Twitter with case of Anti-LGBT campaign in Indonesia using Naïve Bayes, decision tree, and random forest algorithm. *Procedia Computer Science*, *161*, 765–772.

Hastie, T., Tibshirani, R., Friedman, J., & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, *27*(2), 83–85.

Jannah, N. Z. B., & Kusnawi, K. (2024). Comparison of naïve Bayes and SVM in sentiment analysis of product reviews on marketplaces. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, *8*(2), 727–733.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. *European Conference on Machine Learning*, 137–142.

Kang, H., Yoo, S. J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, *39*(5), 6000–6010.

Kaur, N. (2021). SENTIMENT ANALYSIS OF E-BANKING CUSTOMER REVIEWS USING NLP. *ShodhKosh: Journal of Visual*

and Performing Arts, *2*. https://doi.org/10.29121/shodhkosh.v2.i2.2021.5743

Korovkinas, K., Danėnas, P., & Garšva, G. (2017). SVM and Naïve Bayes classification ensemble method for sentiment analysis. *Baltic Journal of Modern Computing.*, *5*(4), 398–409.

Ligthart, A., Catal, C., & Tekinerdogan, B. (2021). Systematic reviews in sentiment analysis: a tertiary study. *Artificial Intelligence Review*, *54*(7), 4997–5053.

Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.

Loria, S. (2018). textblob Documentation. *Release 0.15*, *2*(8), 269.

Martin, J. H., & Jurafsky, D. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (Vol. 23). Pearson/Prentice Hall Upper Saddle River.

McCallum, A. (1998). *A comparison of event models for naive bayes text classification*.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill. https://books.google.co.id/books?id=EoYBngEACAAJ

Monika, P., Kulkarni, C., Kumar, N. H., Shruthi, S., & Vani, V. (2022). Machine learning approaches for sentiment analysis: A survey. *International Journal of Health Sciences*, *6*(S4), 1286–1300.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. *LREc*, *10*(2010), 1320–1326.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135.

Pooja, & Bhalla, R. (2022). A review paper on the role of sentiment analysis in quality education. *SN Computer Science*, *3*(6), 469.

Pradhan, A., & Merline, A. (2023). *Sentiment Analysis: Methods and Application Using Machine Learning in Different Fields*.

Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P.-M. (2023). A review on sentiment analysis from social media platforms. *Expert Systems with Applications*, *223*, 119862.

Sharma, T., & Kaushik, P. (2023). Leveraging sentiment analysis for twitter data to uncover user opinions and emotions. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(8s), 162–169.

Utami, T. W., & Hidayat, A. M. (2024). The effect of e-service quality on customer satisfaction with repurchase intention as a moderating variable in the Soco by Sociolla application. *Edelweiss Applied Science and Technology*, *8*(6), 3886–3901.

Utomo, T. Y., & Hidayat, A. M. (2023). The Intervening Role Of E-Satisfaction In Relationships Between E-Service Quality And E-Loyalty: Case Of BNI Mobile Banking In Indonesia. *Jurnal Aplikasi Manajemen*, *21*(3), 671–687.

Vechtomova, O. (2009). Book review: Introduction to information retrieval by christopher d. manning, prabhakar raghavan, and hinrich schütze. *Computational Linguistics*, *35*(2).

Zhong, N., & Ren, J. (2022). Using sentiment analysis to study the relationship between subjective expression in financial reports and company performance. *Frontiers in Psychology*, *13*, 949881.