Vol. 11, No. 1, April 2026, Page. 243-257
https://journal.uin-alauddin.ac.id/index.php/instek
p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

Sentiment and Trend Analysis of Public Opinion on Bank BTN Mortgage Services via Instagram

Putri Dian Zara[*1], Veny Anggraini[2], Indra Budi[3], Aris Budi Santoso[4], Prabu Kresna Putra[5]

[1,2,34]Faculty of Computer Science, University of Indonesia, Jakarta, Indonesia
[5]Research Center for Data and Information Science, National Research and Innovation Agency, Bandung, Indonesia
Email: [1]putri.dian21@ui.ac.id, [2]veny.anggraini93@gmail.com, [3]indra@ui.ac.id, [4]aris.budi@ui.ac.id, [5]prab003@brin.go.id

Abstrak

Bank BTN merupakan pemain utama layanan Kredit Pemilikan Rumah (KPR) di Indonesia, namun masih menerima banyak keluhan nasabah di media sosial. Berdasarkan Laporan OJK 2023, sektor perbankan mencatat 10.848 pengaduan konsumen, meningkat dari 7.425 pengaduan pada 2022. Kondisi ini menunjukkan perlunya analisis sistematis terhadap opini publik mengenai layanan KPR. Penelitian ini menganalisis sentimen yang diungkapkan melalui akun Instagram resmi BTN selama tahun 2023. Komentar pelanggan diklasifikasikan ke dalam sentimen positif dan negatif menggunakan teknik penambangan teks. Empat algoritma pembelajaran mesin dibandingkan, yaitu K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), dan Naive Bayes (NB). Sebanyak 1.507 komentar diperoleh setelah pembersihan, pra-pemrosesan, dan penyaringan dari total 28.745 komentar. SMOTE diterapkan untuk menangani ketidakseimbangan data. Hasil eksperimen menunjukkan Decision Tree memberikan performa terbaik pada data uji dengan F1-score 0,9348, diikuti oleh SVM (0,9259). Analisis tren juga dilakukan dengan uji korelasi Pearson, Spearman, dan Kendall. Korelasi antara sentimen harian dan jumlah realisasi KPR ternyata lemah dan tidak signifikan secara statistik (Pearson: 0,082, p = 0,187; Spearman: 0,099, p = 0,112; Kendall: 0,077, p = 0,102; N = 259). Hasil ini mengindikasikan bahwa fluktuasi sentimen di Instagram belum dapat dijadikan prediktor langsung bagi pencairan KPR pada hari yang sama. Kemungkinan terdapat jeda waktu, struktur hubungan non-linier, atau pengaruh faktor fundamental yang lebih dominan.

Kata kunci: perbankan, kredit perumahan, instagram, analisis sentimen, media sosial

Abstract

Bank BTN is the leading provider of mortgage (KPR) services in Indonesia, yet customer complaints on social media remain frequent. According to the OJK 2023 Annual Report, the banking sector recorded 10,848 consumer complaints, an increase from 7,425 complaints in 2022. This trend highlights the need for a systematic analysis of public opinion on KPR services. This study analyzes public sentiment expressed on Bank BTN's official Instagram account during 2023. User comments were classified into positive and negative sentiments using text mining techniques. Four machine learning algorithms were compared: K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), and Naive Bayes (NB). A total of 1,507 comments were obtained after data cleaning, preprocessing, and filtering from 28,745 raw comments. SMOTE was applied to address the class imbalance. Decision Tree achieved the highest F1-score (0.9348) on the imbalanced test set, followed by SVM (0.9259). Trend analysis was conducted using Pearson, Spearman, and Kendall correlation tests. The correlation between daily sentiment and KPR disbursement was weak and not statistically significant (Pearson: 0.082, p = 0.187; Spearman: 0.099, p = 0.112; Kendall: 0.077, p = 0.102; N = 259). These results indicate that sentiment fluctuations on Instagram cannot be used as a direct predictor of same-day mortgage disbursement. There may be a time lag, a non-linear relationship structure, or the influence of more dominant fundamental factors.

Keywords: *banking, mortgage, instagram, sentiment analysis, social media*

This work is an open access article and licensed under a Creative Commons Attribution-NonCommercial ShareAlike 4.0 International (CC BY-NC-SA 4.0)

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

1. INTRODUCTION

The internet has become a core pillar of the modern information society, driving the rise of social media platforms that have changed the way people communicate and share information [1]. Many platforms have increased the value of customer feedback exchanged. Social media has become a platform for customers to express their dissatisfaction and complaints regarding banking services [2]. The Indonesian Consumer Protection Agency (YLKI) reports that most consumer complaints are related to banking services. The Financial Services Authority (OJK), Indonesia's financial services authority, has recognized the importance of social media as a source of real-time public feedback, including service complaints [3].

Numerous studies have analyzed customer feedback on social media using text categorization. One study focused on sentiment analysis and identified consumer complaints in user posts. Other methodologies use sentiment analysis to identify posts containing complaints about products or services [4]. Another study identified key themes and topics related to customer complaints and used this information to improve service. However, previous studies have limitations, including ignoring complaints from both positive and neutral subjects [5].

The sentiment analysis model was also used in this case to identify complaints from users of the OVO and DANA digital wallet services by comparing the following algorithms: K-Nearest Neighbor (KNN), Naive Bayes, Decision Tree, and Random Forest. This study used 5,000 Twitter reviews, with a composition of 20% positive and 80% negative. Experimental results showed that KNN achieved the highest accuracy of 91.8% [6]. Sentiment analysis was also applied to assess customer satisfaction on mobile banking platforms at Indonesia's largest banks, such as BCA Mobile, Livin' by Mandiri, BRI Mobile (Brimo), and BNI Mobile. Sentiment data was collected from Twitter and then used the Support Vector Machine (SVM) algorithm with an accuracy of 92.5%, and the most positive sentiment results were for Brimo [7].

Bank BTN is the first state-owned bank appointed by Indonesian Government to distribute mortgage services to the people, the mortgage product commonly known as "KPR" in Indonesia. At the age of 75 this year, BTN is still the main player in disbursing home loans. In addition, BTN is the only bank with the largest mortgage contribution, reaching 47% of total BTN credit as of December 2023 [8]. Despite BTN's dominant position as "the main player housing mortgage" in Indonesia, there are still quite a lot of complaints on social media. According to the OJK Annual Report 2023, the banking sector recorded the highest number of consumer complaints among all financial service sectors in Indonesia, with 10,848 complaints received through the OJK Consumer Protection Portal (APPK) throughout 2023, representing the largest share across all sectors [9]. This figure reflects a substantial increase from 7,425 banking complaints recorded in 2022 [10], indicating a growing trend of consumer dissatisfaction with banking services. Furthermore, the mediakonsumen.com platform which shows that there are 68 customer complaints regarding mortgage services based on query search "KPR BTN" [11]. This highlights the need to analyze customer sentiments on BTN's official Instagram account where direct communication can occur between Bank and customers.

Instagram has become a popular social media platform in Indonesia. Instagram users generally express their opinions, experiences, and concerns, making it a valuable resource for understanding customer sentiments and preferences in real-time, which is crucial for businesses seeking to improve their services and offerings [12]. Moreover, BTN has a verified account and specifically manages this media platform. This study therefore aims to: (1) identify customer sentiment on Bank BTN's official Instagram account regarding KPR services, and (2) analyze the correlation between customer sentiment and KPR disbursement trends.

This study differs from previous studies in several key aspects: first, it focuses specifically on mortgage loan (KPR) services rather than general banking services; second, it uses data derived Instagram, which currently serves as BTN's most interactive social media channel for customer engagement. And third, it includes an analysis of mortgage loan realization trends. This study is organized into five sections. After the introduction, the Literature Review section discusses previous related research, the Methodology section explains the data collection and analysis techniques, the Analysis and Result section presents the findings, and the Conclusion section summarizes the study and offers recommendations for future research.

2. RESEARCH METHODOLOGY

In this section, researchers provide several methods to perform this research including Data Collection, Data Preprocessing, Vectorization, Sentiment Analysis, and Evaluation Model Performance.

2.1 Data Collection

The process of gathering customer feedback data involved scraping data from the Instagram platform using Python libraries. The feedback data was obtained using the Python library such as "instaloader". Furthermore, data was extracted from posts on the official BTN Instagram account between 1 January 2023 and 31 December 2023 along with all comments on these posts. The time segmentation was determined because the distribution of Nonsubsidized segment by Bank BTN has increased in terms of value and volume.

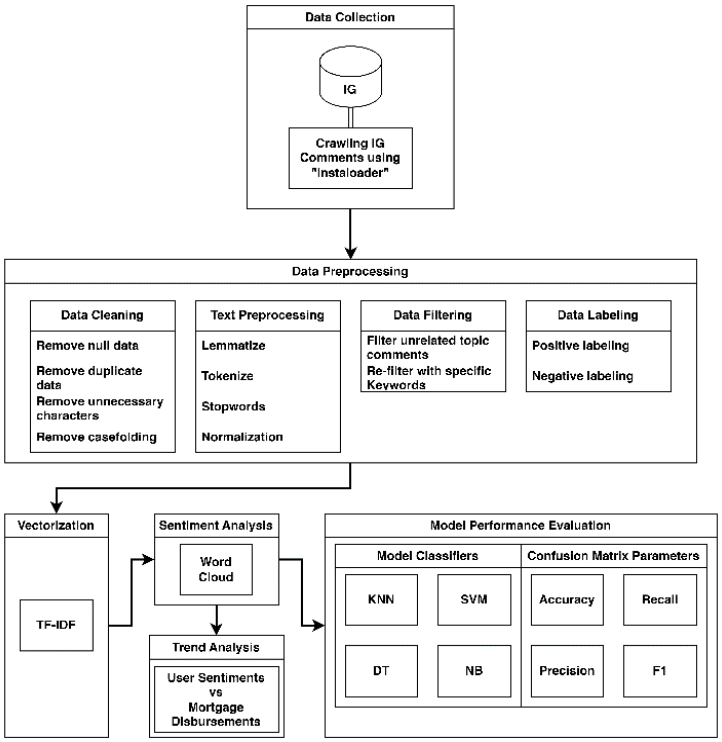


Figure 1. Process Flow

2.2 Data Preprocessing

During the preprocessing of data, various procedures was implemented to convert raw data into refined data. Refer to researchers [7] and [13], these steps include data cleansing, case folding, tokenization, normalization, stop word removal, and stemming. Here's a breakdown of each step:

1. Data Cleaning

Data cleaning is performed to remove null data, duplicate data, and redundant or irrelevant content, such as URLs, mentions, emojis, or unknown characters. Furthermore, this step eliminates case folding, which involves converting all letters to lowercase. Case folding ensures that each word in a comment, once converted to lowercase, aligns with the word index dataset for further processing. This entire process is crucial for reducing interference with machine learning algorithms and handling smaller text data, thereby improving information accuracy.

2. Lemmatize

Lemmatize is performed by using lemmatize function in NLP-ID library. This choice is motivated by the Indonesian language used in the dataset. The lemmatization process, accessed through the Lemmatizer class in the NLP-ID library, handles prefixes like "me-", "di-", "ke-", "ter-", and "ber-", as well as suffixes like "-an" and is used to get the root words from every word in a sentence. By converting words to their base forms, lemmatize ensures that the model encounters more consistent vocabulary.

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

3. Tokenization

Tokenization involves breaking text into smaller units, typically words or subwords, which is called as tokens. This process helps understand the structure of text and makes it easier to analyze or process. Tokens can be individual words, punctuation marks, or even sub words got through techniques like word segmentation or sub word tokenization. For example, tokenizing the sentence "The cat is sleeping." would result in tokens: ["The", "cat", "is", "sleeping", "."].

4. Stopwords

Stop words, identified using the NLP, Sastrawi Python and other library, that removed from the text. These stop words, were taken from the StopWordRemoverFactory class in the Sastrawi library, include common terms like "dan", "di", "dari", "para", "yang", "namun", "jika", "lalu", "sehingga", "kemudian", and "oleh karena". For each comment that is retained at this stage, only for words that have important meaning.

5. Normalization

Normalization refers to the process of transforming text into a standard format. As several words' structure might be in slang words, normalization should be done to normalize these slang words to the standard format. For example, words such as "gue", "gw", "guwe", and "gua" should be normalized become "saya" as the formal word.

6. Filtering

Data filtering is conducted to get related comments with the topic of this resarch subjects. As many comments have diverse subject topics, it might be not coherent with the objectives of the research. Therefore, data filtering should be done to highlights specific information related to the study. The 25 keywords on Table 1 were selected through a combination of term-frequency analysis on raw data, expert judgement from KPR practitioners at Bank BTN, and validation against banking and property terminology dictionaries, covering four thematic categories: financial aspects, application processes, KPR products, and ecosystem stakeholders.

7. Labeling

In the final step of data preprocessing, this study conduct data labeling based on cleaned data. The labeling process can be done manually or using several libraries to analyze all user comments. The comments are categorized as positive label if several positive words and words understanding have positive meaning. It also happens for negative label when several negative words and words understanding have negative meaning. Afterward, these labels can be used to create modeling for sentiment analysis.

2.3 Vectorization

Vectorization will be employed to transform a document or text into a matrix, facilitating its processing through mathematical methods. This research uses one of the largest vectorization methods, namely Term Frequency-Inverse Data Frequency or TF-IDF. The TF-IDF implementation stage will count each word in the document as a feature.

2.4 Sentiment Analysis

Sentiment analysis identifies and extracts subjective information, such as opinions or attitudes, from language, and determines if the sentiment is positive, negative or neutral [7]. Indonesian language will be used during analysis to categorize the sentiments. Therefore, the result of sentiment analysis are values that represent positive or negative opinions. Understanding the characteristics of these sentiments can be helpful in designing strategies and making the right decisions to enhance products or services [7].

In this study, sentiment analysis will be conducted using popular classifiers such as K-Nearest Neighbors, Support Vector Machine, Decision Tree and Naive Bayes. All classification models are evaluated for their accuracy, precision, recall, and F1-Score by perform K-fold cross-validation technique. Cross-validation is a technique that divides the original dataset into several subsets. This allows the training data that was initially processed in one section will be rotated across a few K subsets.

The test data is kept aside for later use to confirm and validate the preliminary analysis. This study will use stratified 10-fold cross-validation because it is one of the recommended cross-validations for selecting the best model. The stratified strategy preserves the original class proportion (positive vs.

negative) within each fold, which is essential for handling the imbalanced dataset and ensuring reliable performance estimates across all folds [13]. Two kinds of classification that will be used in this study are positive and negative sentiment from Instagram users comment toward KPR Bank BTN services.

Considering the potential class imbalance in the collected dataset, class imbalance represents a fundamental challenge in sentiment analysis, as the underrepresentation of minority classes can bias the classifier toward the majority class and reduce its effectiveness in detecting minority-class instances[14]. The Synthetic Minority Over-sampling Technique (SMOTE), introduced by Chawla et al. (2002), mitigates class imbalance by generating synthetic samples for the minority class through linear interpolation between a given instance and one of its *k*-nearest neighbors. In contrast to simple random replication, this approach increases sample diversity, thereby reducing the risk of overfitting and enhancing the model's ability to generalize to minority-class instances. Recent evaluations by Taskiran et al. (2025) demonstrate that SMOTE-based oversampling can substantially improve classification fairness in imbalanced textual datasets, while also highlighting that such techniques should be applied solely to the training data to avoid data leakage.

- K Nearest Neighbors (kNN)

K-nearest neighbors classification is performed by identifying the K closest documents from the training data using a distance measure (Jiang, Pang, Wu, & Kuang, 2012). The new document is then assigned to the class that is most common among these K closest documents. To avoid ties, K is typically chosen as an odd number [15]. The method of K-nearest neighbor works well when the amount of training data is large [15].

- Support Vector Machine (SVM)

The SVM algorithm is a machine learning classification technique that transforms a set of data points into a new space using a kernel function, especially when the original data is not linearly separable [16]. The SVM algorithm used to classify text using index weights as features [13].

- Decision Tree (DT)

One of the earliest and commonly used classifiers, decision trees (DT) are built on an intuitive method of learning through a hierarchical tree structure that breaks down the data space. The goal is to establish a tree where each internal node is accountable for analyzing a feature and steering unlabeled instances towards one of its children based on the feature's value. While leaf nodes are representing specific categories. Even though they are intuitive, the fundamental design of decision trees can be vulnerable to minor data fluctuations and overfitting problems [17].

- Naive Bayes (NB)

Naive Bayes (NB) is a probabilistic algorithm in machine learning based on Bayes Theorem, which is used in classification tasks [18]. In mathematics, Bayes theorem is used to compute conditional probabilities, assuming features are conditionally independent and of equal importance. It has also been shown that Naive Bayes works accurately in identifying the actual polarity of the given sentence, even when dealing with small, imbalanced datasets [19].

2.5 Model Performance Evaluation

Model performance evaluation serves to assess model classification. Model performance evaluation is necessary because it allows us to determine the model's performance and obtain the level of precision of data predictions [7]. Predictions can be true positive, true negative, false positive, or false negative. A true positive is defined as the predicted positive value matching the actual positive value. Conversely, a false positive arises when a predicted positive value does not align with the actual negative value. A true negative happens when the predicted negative value matches the actual negative value, while a false negative happens when the predicted negative value does not match the actual positive value. These four parameters are contained in the confusion matrix.

Based on confusion matrix, three measurements can be conducted to find scoring of model performance evaluation such as accuracy, precision, and recall. Accuracy represents the proportion of accurate predictions among all prediction data. Precision shows the relationship among correct positive. predictions and all positive predictions made. Similarly, recall evaluates the proportion of accurate positive predictions to all potential true positive values, including both true positives and false negatives.

After completing the classification modeling using training and testing data for each algorithm and displaying the prediction model output in the form of a confusion matrix, the next step is to present the

evaluation metric calculations in a classification report. The F1-score is the most appropriate evaluation metric because it provides a more objective picture of performance through a weighted average of precision and recall. By combining these two aspects, the F1-score effectively considers the impact of prediction errors in the form of false positives and false negatives [20].

2.6 Trend Analysis

Trend analysis can identify patterns, shifts, and trajectories within the overall public sentiment, which is crucial to understanding the evolution of public perception and the potential impact on consumer behavior or brand reputation [21]. This methodology surpasses statistical summaries, provide insights that can be communicated in the institution's strategic decision-making process [22].

To systematically evaluate trends and correlate them with a variety of other factors (e.g., service uptake, market efficacy, or unique events), statistical corr

elation frameworks such as Pearson, Spearman, and Kendall are crucial. These can utilize statistical coefficients that measure the strength and orientation of the relationship from multiple variables. For example, the Spearman Rank correlation test has been used in similar studies to confirm the correlation between sentiment and financial indicators, even in case where the correlation is minimal [23].

While Pearson's coefficient measures a linear relationships, Spearman and Kendall serve as non-parametric alternatives that are very precise for non-normally distributed datasets or ordinal scales, which are often used in sentiment evaluations or rankings. The integration of these diverse correlation methodologies facilitates a holistic comprehension of the ways in which public opinion trends on social media and mass consumption behavior [22].

3. RESULTS AND DISCUSSION

This section will explain the process of each research stage, including data processing, the analysis and results of this study.

3.1 Data Description

This study collected user comments from every BTN post in period 1 January 2023 until 31 December 2023. The total crawl data is 28,745 user comments during the period by collecting its metadata from 170 posts on BTN’s official Instagram account. The data stored in 8 columns such as post id, post caption, postdate created, comment id, comment username, comment content, and comment date created.

Table 1. Keywords Filter

Keywords
"jadwal", "jangka waktu", "subsidi", "akad", "take over", "lunas", "promo", "simulasi", "pelayanan", "pengajuan", "perumahan", "tunggakan", "denda", "developer", "property", "properti", "proses", "syarat", "dokumen", "biaya", "dp", "bunga", "rate", "hgb", "sertifikat", "shm"

In the data preprocessing stage, null data was not found but duplicated comment id was found in 4,202 data and duplicated comment’s contents with the same username was found in 598 data. As a result, the remaining comments are 23,945 data.

Next step is data cleaning to remove unnecessary texts (mentions, hashtags, numbers, emojis, incomplete URL, multiple spaces, punctuation, single char, and repeated chars) also removing case folding to reduce interference and enhance the accuracy of information for machine learning modeling processes.

Additionally, text preprocessing was conducted such as lemmatizing, tokenizing, removing stopwords, and text normalization. This process is important to eliminate words that have no meaning

with related topics. This process uses several open-source Python libraries such as NLTK, NLP-ID by Kumparan, Sastrawi, and other libraries from open-source GitHub repositories.

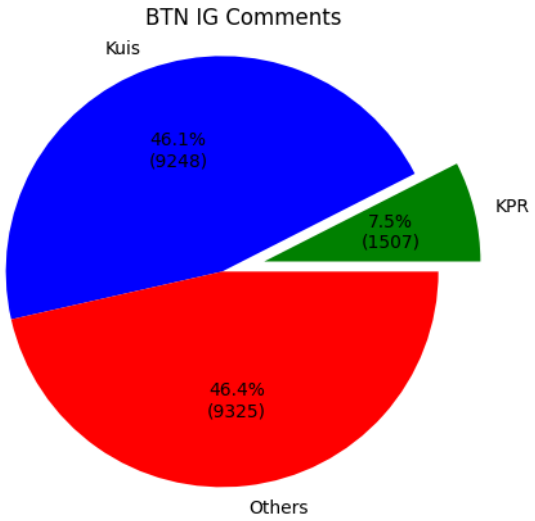


Figure 2. Topic Categories Data

As the subject in this study is focusing on housing mortgages, the next step is conducting filter comments related to KPR in Bank BTN. To do the filter, a set of keyword-based filters was applied, as presented in Table 1, to reduce noise from irrelevant data. This way, the dataset would reflect public opinion focused on KPR processes, experiences, and constraints. While analyzing the data, around 62.5% of comments are related to quiz events, specifically promotional quiz campaigns organized by Bank BTN that are unrelated to KPR services (**Error! Reference source not found.**). Therefore, comments associated with quiz tags were also filtered out. As a result, the total comments of housing mortgage are 1,507 data.

The final step is creating a label for each comment such as positive and negative label. The labeling process was conducted manually in collaboration with mortgage experts by considering the overall meaning of each comment, as shown in Table 2. The results indicate that 1,333 comments were labeled as positive and 174 as negative (**Error! Reference source not found.**), this indicates that the dataset is imbalance.

Table 2. Example Dataset Labeling

No	Comment	Label
1	Min mau nanya, bank BTN ada kartu kreditnya ga...	Positive
2	Wow,, banyak promo	Positive
3	Jangan hanya komersil saja, pastikan tanggung ...	Negative
4	Bank BTN sangat MENGECEWAKAN utk bagian team A...	Negative
5	Makin lama bunganya tinggi banget	Negative

3.2 Vectorization

The vectorization step was conducted using TF-IDF that is performed to get the feature extraction and count of words based on its frequency and importance in a feature. In this study, the scikit-learn library was utilized for word feature extraction and sort them based on the number of times the words appear. The higher TF-IDF score means less frequent terms in documents and vice versa. For example,

a lower score (e.g., "pakai" (use) and "ambil" (take)) means the more common the word across documents.

Table 3 display the top 20 most important terms from data. It achieves this by taking the ranking DataFrame, which contains each term and its aggregated TF-IDF score (represented as 'rank'), sorting this DataFrame in descending order based on the 'rank' column, and then selecting the first 20 entries. This effectively shows the terms that have the highest overall importance or frequency across the data, according to the TF-IDF calculation. The result Table 3 showed the most user comments in KPR subjects are related to mortgage administration, bill payments, mortgage subsidies, and promotions. In contrast, the more unique words in the first document (e.g., "kpr" and "rumah"), the higher the score is. The word "KPR" is among the top 10 the most frequently words occurring in customer comments. The result of vectorization will be used to adjust the classifier model algorithm for prediction.

Table 3. Most Frequently Words

No	Term	Rank	No	Term	Rank
1	promo (promotion)	124,770,100	6	bayar (pay)	60,008,769
2	admin (administration)	84,949,596	7	rumah (house)	59,820,121
3	kpr (mortgage)	78,857,402	8	provisi (provision)	53,795,882
4	subsidi (subsidized)	78,454,625	9	pakai (use)	53,414,002
5	biaya (fee)	69,281,871	10	ambil (take)	48,471.445

Out of 1,507 user comments related to KPR subjects, 1,333 comments have positive sentiment, and 174 comments have negative sentiment. To view the most frequently appearing words in user comments, we used word cloud for each sentiment label on Figure 3. The green word cloud state positive sentiment, and the red word cloud state negative sentiment. This visual representation helps identify terms presented in the data set, and it is possible to understand the main topics and expressed emotions representing the set of comments. In the positive word cloud, terms such as "kpr", "promo", and "rumah" are dominant, reflecting discussions about promotional offers and the mortgage process itself. In contrast, the negative word cloud is dominated by terms such as "kecewa" (disappointed), "bunga" (interest rate), and "lama" (slow/long), indicating customer dissatisfaction with high interest rates and slow processing.

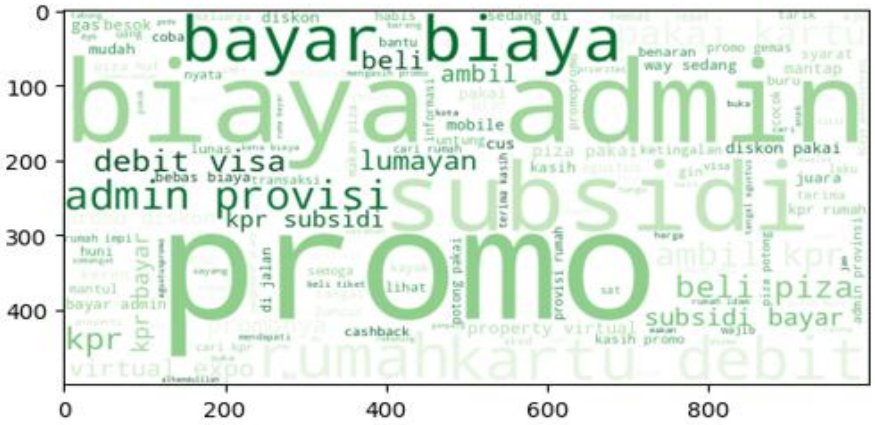
Due to the imbalanced dataset in this study, SMOTE was applied after TF-IDF vectorization and before sentiment classification to address the imbalanced data in the training phase. To balance the classes, oversampling the minority class was performed, ensuring that the sentiment classification model detects both positive and negative sentiments. This improves generalization and reduces biased predictions that favor the majority class. SMOTE was implemented using the imbalanced-learn library with k_neighbors = 5, sampling_strategy = 1.0 (target ratio 1:1), and random_state = 42.

In the subsequent step, machine learning modeling was conducted using four algorithms, namely Decision Tree (DT), K-Nearest Neighbor (KNN), Naive Bayes (NB), and Support Vector Machine (SVM), to classify sentiment categories. User comment data were utilized as the dataset for both training and testing processes. This study employed the "PyCaret" library to perform machine learning modeling by configuring several experimental settings, including the target variable, train–test split ratio, and stratified K-fold cross-validation with 10 folds (fold_strategy='stratifiedkfold' in PyCaret), which preserves the original class distribution within each fold. All four classifiers were implemented using the PyCaret 3.x classification module with default hyperparameters to ensure a fair baseline comparison without introducing classifier-specific tuning bias.

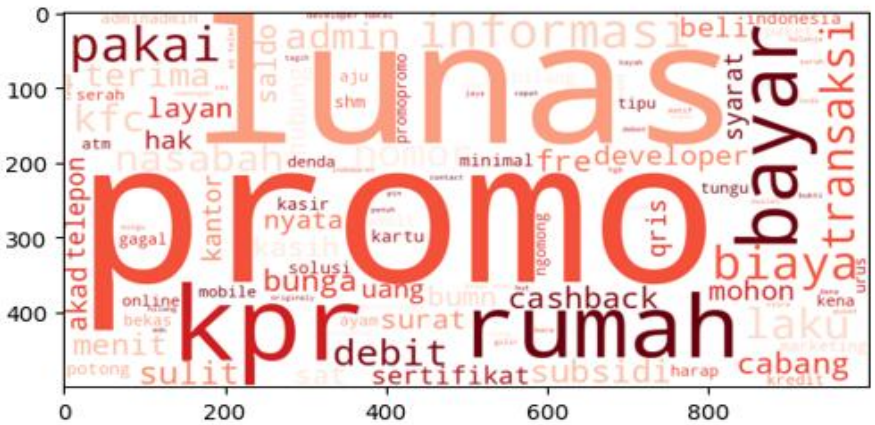
The KNN model was implemented using scikit-learn's KNeighborsClassifier with k=5 neighbors and uniform weighting based on Euclidean distance. The SVM model was implemented as a linear Support Vector Machine through SGDClassifier (PyCaret's default implementation for the svm identifier) using hinge loss, L2 regularization, and α=0.0001. The Decision Tree model employed the Gini impurity criterion without a maximum depth constraint, allowing the tree to grow until all terminal

nodes were pure or contained fewer samples than the minimum split requirement. Meanwhile, the Naive Bayes classifier utilized GaussianNB with the default variance smoothing parameter (var_smoothing = 1e^{-9}), which is suitable for continuous TF-IDF features. To ensure reproducibility, a fixed random_state = 42 (session_id in PyCaret) was applied consistently across all experiments.

First, the label was defined as the target variable. Second, the dataset was split using the default ratio provided by PyCaret, with 70% of the data allocated for training and 30% for testing. A total of 1,054 data were used for training and 453 data for testing. The original data imbalance ratio between positive and negative (1,333:174) was 7.66:1. SMOTE was applied only to the training set after the train-test split, while the test set retained its original imbalanced distribution to prevent data leakage and ensure realistic evaluation. The training data was then split with a 70:30 ratio, resulting in 933 positive and 121 negative samples. SMOTE was then performed on the training data, resulting in a 1:1 ratio. This resulted in 932 samples for each positive and negative sample, for a total of 1,864 samples, and 453 samples for testing (fixed).



a. positive



b. negative

Figure 3. Word cloud on positive and negative reviews

Before applying SMOTE, all models showed high recall values, indicating a tendency for the models to recognize the dominant class over the majority. The DT model recorded the highest F1-score at 0.9443, followed by KNN (0.9402) and SVM (0.9412). However, the high recall (≥0.94) indicates a bias towards the positive class due to the synchronization of data distribution. The NB model had the lowest accuracy (0.7881) despite relatively high precision, indicating that the model tends to be cautious in classifying the minority class.

Vol. 11, No. 1, April 2026, Page. 243-257
https://journal.uin-alauddin.ac.id/index.php/instek
p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

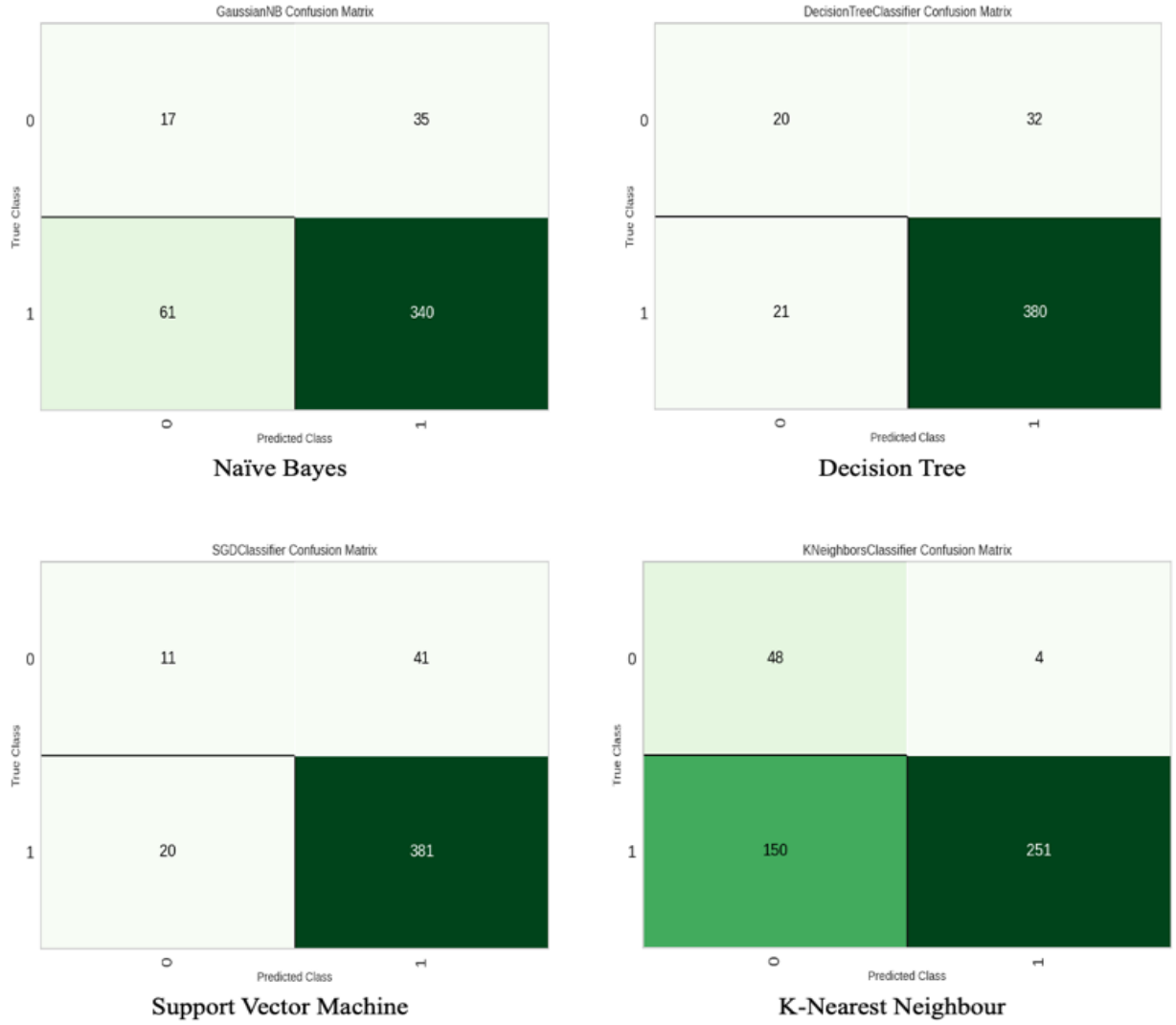


Figure 4. Confusion Matrix After SMOTE

From the Decision Tree confusion matrix (Figure 4) that is True Positive (TP) = 380 is predict positive sentiment correctly, True Negative (TN) = 20 is correctly predict negative sentiments. False Positive (FP) = 32 shows Negative sentiments incorrectly identified as positive and False Negative (FN) = 21 shows Positive sentiments incorrectly predicted as negative. The low True Negative count (20) relative to False Positive (32) indicates the model struggles to identify negative sentiments, achieving only 21,15% recall on the minority class. SMOTE only applied on training data for modeling, but not on test data. This highlights the challenge of training on balanced data (with SMOTE applied) while testing on imbalanced data.

After applying SMOTE, Decision Tree retained the highest overall F1-score after SMOTE at 0.9348, with recall of 0.9476 and precision of 0.9223, indicating a well-balanced precision–recall trade-off even when evaluated on the original imbalanced test set. The relatively small drop from its pre-SMOTE F1 (0.9443) suggests that DT is robust to distributional changes introduced by synthetic oversampling, while still improving its ability to detect negative sentiment instances.

Naive Bayes showed no change in performance after SMOTE, maintaining an F1-score of 0.8763, recall of 0.8479, and precision of 0.9067. This stability suggests that the Gaussian assumption underlying the GaussianNB implementation limits its sensitivity to changes in training data distribution, making SMOTE neither beneficial nor detrimental for this classifier.

Table 4. Performance Measures Classification Model

Before SMOTE

	Model	Accuracy	Recall	Prec.	F1
1	DT	0,8985	0,9726	0,9176	0,9443
2	KNN	0,8896	0,9800	0,9034	0,9402
3	SVM	0,8896	0,9975	0,8909	0,9412
4	NB	0,7881	0,8479	0,9067	0,8763

After SMOTE

	Model	Accuracy	Recall	Prec.	F1
1	DT	0,883	0,9476	0,9223	0,9348
2	SVM	0,8653	0,9501	0,9028	0,9259
3	NB	0,7881	0,8479	0,9067	0,8763
4	KNN	0,66	0,6259	0,9843	0,7652

The result of sentiment analysis modeling (Table 4) has four score metrics such as accuracy, recall, precision, and F1. The highest score indicates overall effectiveness in correctly classifying instances for classifying the sentiment. In addition, the result shows that the best performing classifier differs between the two scenarios. Before applying SMOTE, DT achieved the highest F1-score, this apparent superiority is largely driven by the dominant positive class in the imbalanced training data. After applying SMOTE and evaluating on the imbalanced test set, DT still achieved the highest F1-score, indicating that DT provides the most balanced precision–recall trade-off when the model is forced to genuinely discriminate the minority class.

The experimental results demonstrate the critical importance of appropriate metric selection and data balancing techniques when evaluating classifiers on imbalanced datasets. Before applying SMOTE, the models exhibited high accuracy (78.81-89.95%) and recall values, yet the confusion matrices reveal significant class imbalance effects, particularly evident in the K-Nearest Neighbour classifier where 150 false negatives indicates a critical weakness in detecting the minority class (negative sentiment). Thus K-Nearest Neighbour model shows a significant lack of sensitivity to customer complaints.

3.3 Trend Analysis

Based on data from BTN analyst meeting presentation, it is known that there was a significant increase in mortgage disbursement growth especially from subsidized and non-subsidized mortgage from Q1 to Q3 in 2023, with the greatest growth occurring between Q2 and Q3, but there was also a slight decline from Q3 to Q4 [24],[25],[26],[27],[28].

To conduct trend analysis, statistical correlation analysis was performed using Pearson correlation, Spearman correlation, and Kendall rank correlation. At first, seasonal decomposition is performed to identify the trend, seasonal, and residual components of the collected time series as shown (Figure 5). The STL decomposition of daily mortgage data shows a clear weekly seasonal pattern, with loan activity only occurring on weekdays. The trend component shows an upward movement in the first half of 2023. Random fluctuations (residuals) remain quite high, indicating the presence of external factors or daily variability not yet explained by seasonal and trend patterns.

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

The results of the statistical correlation test (The practical implication is that, in order to exploit social media sentiment as an early indicator of KPR disbursement trends, more sophisticated analytical approaches are needed, including Granger causality tests, autoregressive time-series models with lagged sentiment inputs, and data aggregation at weekly or monthly resolutions to reduce daily noise and enable the detection of medium-term relationship patterns.

Based on **Error! Not a valid bookmark self-reference.**, these findings indicate that daily fluctuations in Instagram sentiment cannot yet serve as direct predictors of same-day KPR disbursement. Several external factors may explain this weak relationship. The consistency across all three methods in producing low and non-significant coefficients reinforces the conclusion that, at a daily resolution, there is insufficient statistical evidence to support the existence of either a linear or a monotonic relationship between public sentiment and KPR disbursement.

Table 5), shows that the relationship between daily sentiment and KPR disbursement amounts is consistently weak across all three measurement methods. The Pearson correlation coefficient of 0.082 ($p = 0.187$), the Spearman rank correlation coefficient of 0.099 ($p = 0.112$), and the Kendall rank correlation coefficient of 0.077 ($p = 0.102$) are all non significant at the $\alpha = 0.05$ level. The consistency between the parametric method (Pearson) and the non parametric methods (Spearman, Kendall) in producing low and non significant coefficients reinforces the conclusion that, at a daily resolution, there is insufficient statistical evidence to support either a linear or a monotonic relationship between Instagram public sentiment and Bank BTN's KPR disbursement.

The sample size ($N = 259$) refers to the number of working days included in the analysis. This is because the configuration WEEKDAYS_ONLY = True; instructed the script to include only data from Monday to Friday, resulting in 259 business days within the observed period. This result further confirms that the low observed coefficients are not an artifact of any particular distributional assumption, but rather reflect a genuinely weak underlying relationship at daily granularity.

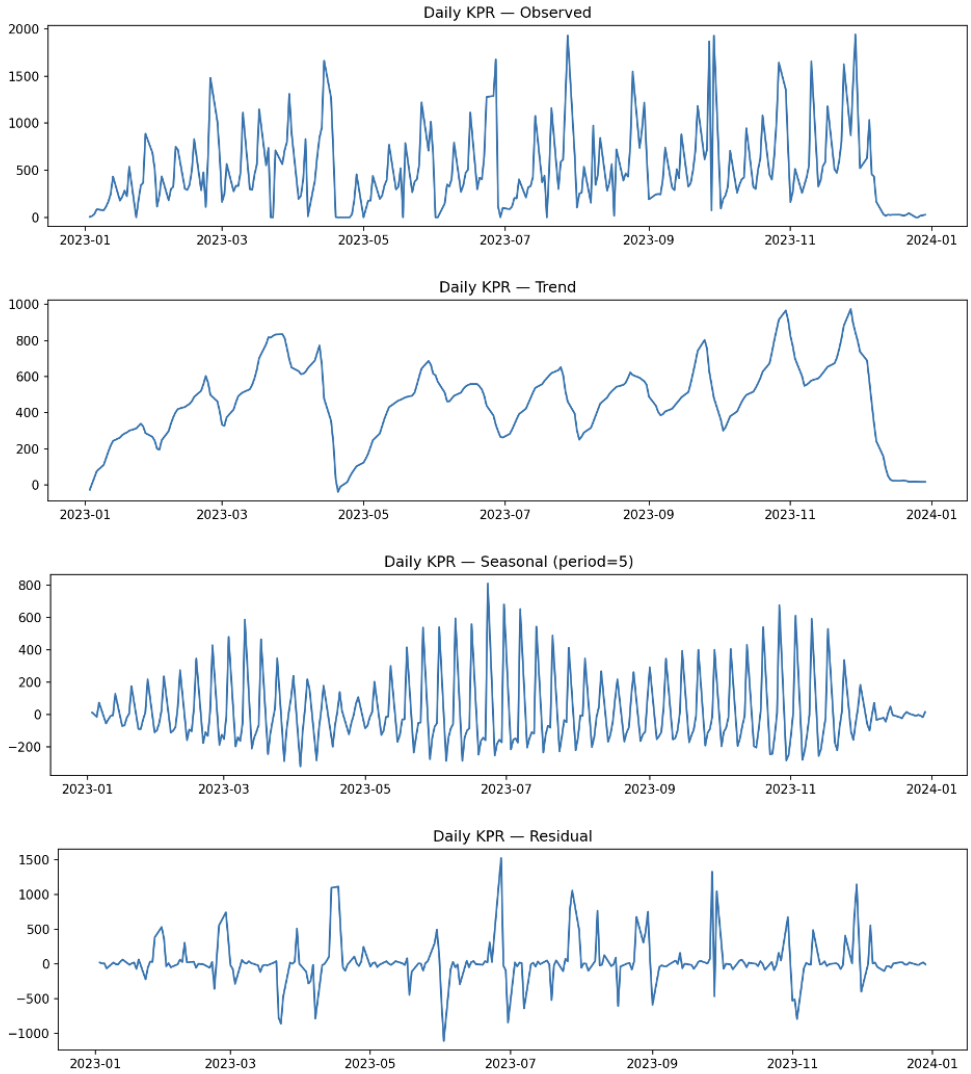


Figure 5. Seasonal Decomposition

The practical implication is that, in order to exploit social media sentiment as an early indicator of KPR disbursement trends, more sophisticated analytical approaches are needed, including Granger causality tests, autoregressive time-series models with lagged sentiment inputs, and data aggregation at weekly or monthly resolutions to reduce daily noise and enable the detection of medium-term relationship patterns.

Based on **Error! Not a valid bookmark self-reference.**, these findings indicate that daily fluctuations in Instagram sentiment cannot yet serve as direct predictors of same-day KPR disbursement. Several external factors may explain this weak relationship. The consistency across all three methods in producing low and non-significant coefficients reinforces the conclusion that, at a daily resolution, there is insufficient statistical evidence to support the existence of either a linear or a monotonic relationship between public sentiment and KPR disbursement.

Table 5. Statistical Data Results

freq	method	coefficient	p_value	N
Daily	Pearson	0.0823	0.1866	259
Daily	Spearman	0.0989	0.1122	259
Daily	Kendall	0.0774	0.1019	259

4. CONCLUSION

This study analyzes 1,507 relevant comments during 2023, customer sentiment on Bank BTN's official Instagram account regarding KPR services, most of them contains positive sentiment (88,5% positive vs. 11,5% negative). The most frequently discussed topics include mortgage administration procedures, bill payments, subsidized mortgage programs, and promotional offers. Modeling prediction is carried out using four algorithms, including Naïve Bayes, Decision Tree, K-Nearest Neighbor, and Support Vector Machine. The application of SMOTE successfully reduced bias toward the majority class and improved performance balance between classes, with the DT model showing the best overall performance on the imbalanced test set, closely followed by SVM. The SVM model achieved the highest recall on the positive class, while KNN showed the highest precision with the lower recall, illustrating different precision recall trade-offs across algorithms on imbalanced data.

The statistical correlation analysis reveals that the relationship between daily public sentiment on Instagram and Bank BTN's KPR disbursement amounts is weak and not statistically significant across all three methods examined. This indicates that, the absence of a significant relationship is not an artifact of any particular statistical assumption, but rather a reflection of the complex array of factors influencing KPR disbursement decisions. Although the statistical relationship was found to be weak, the proposed approach establishes a preliminary framework for future research employing more advanced techniques, such as Granger causality analysis, time-series forecasting, and multi-level temporal aggregation.

Analysis was limited to Bank BTN's official Instagram account only. Customer sentiments expressed on other platforms for through offline channels were not captured, potentially introducing selection bias toward digitally engaged customers. In this study, SMOTE successfully improved training performance, synthetic minority class samples may not fully represent the diversity of real negative sentiments, potentially limiting model robustness for real-world applications. The correlation analysis establishes association but cannot determine causality (whether sentiment drives disbursement or vice versa), requiring more sophisticated econometric approaches for causal claims.

Suggestions for further research are to explore alternative balancing techniques because SMOTE has limitations for real-world development, expand data sources from other social media and collect data over a longer time period, conduct more in-depth trend analysis and qualitative analysis also try to benchmark against other Indonesian Bank with any specific topic.

5. REFERENCE

[1] Statista.com, "Facebook - statistics & facts." Accessed: Oct. 20, 2025. [Online]. Available: https://www.statista.com/topics/751/facebook/?srsltid=AfmBOop5ASpqIdHUGdjwcF7wsZ9wOJHyne1Dysdttb8BVRWvT3HmT7l1#topicOverview

[2] V. Sigurdsson, N. M. Larsen, H. K. Gudmundsdottir, M. H. Alemu, R. G. V. Menon, and A. Fagerstrøm, "Social media: Where customers air their troubles—How to respond to them?," *Journal of Innovation and Knowledge*, vol. 6, no. 4, pp. 257–267, Oct. 2021, doi: 10.1016/j.jik.2021.07.001.

[3] J. Setiawan, A. Milenia, and A. Faza, "An Integrated Approach for Sentiment Analysis and Topic Modeling of a Digital Bank in Indonesia using Naïve Bayes and Latent Dirichlet Allocation Algorithms on Social Media Data," in *2023 4th International Conference on Big Data Analytics and Practices, IBDAP 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/IBDAP58581.2023.10271956.

[4] J. Wang, J. Y. Lai, and Y. H. Lin, "Social media analytics for mining customer complaints to explore product opportunities," *Comput. Ind. Eng.*, vol. 178, Apr. 2023, doi: 10.1016/j.cie.2023.109104.

[5] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning-Based Text Classification," Jun. 01, 2021, *Association for Computing Machinery*. doi: 10.1145/3439726.

[6] I. Tahyudin, A. R. Hananto, S. A. Rahayu, R. M. Anjani, and A. Nurhopipah, "Sentiment Analysis Model Development on E-Money Service Complaints," *TEM Journal*, vol. 12, no. 4, pp. 2050–2055, Nov. 2023, doi: 10.18421/TEM124-15.

[7] M. A. Ligiarta and Y. Ruldeviyani, "Customer Satisfaction Analysis of Mobile Banking Application Based on Twitter Data," in *Proceedings - 2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System, ICE3IS 2022*, Institute of

Electrical and Electronics Engineers Inc., 2022, pp. 322–327. doi: 10.1109/ICE3IS56585.2022.10010221.

[8] N. Dwiantika, "BTN Jadi raja KPR di Indonesia," https://keuangan.kontan.co.id/news/btn-jadi-raja-kpr-di-indonesia. Accessed: Apr. 02, 2024. [Online]. Available: https://keuangan.kontan.co.id/news/btn-jadi-raja-kpr-di-indonesia

[9] "Laporan Tahunan OJK 2023".

[10] "Laporan Tahunan OJK 2022".

[11] Media Konsumen, "Media Konsumen." Accessed: Apr. 01, 2024. [Online]. Available: https://mediakonsumen.com/?s=kpr+btn#gsc.tab=0

[12] H. Sudira, A. L. Diar, and Y. Ruldeviyani, "Instagram Sentiment Analysis with Naive Bayes and KNN: Exploring Customer Satisfaction of Digital Payment Services in Indonesia," 2019.

[13] Y. Handayani, A. R. Hakim, and Muljono, "Sentiment Analysis of Bank BNI User Comments Using the Support Vector Machine Method," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep. 2020, pp. 202–207. doi: 10.1109/iSemantic50169.2020.9234230.

[14] S. F. Taskiran, B. Turkoglu, E. Kaya, and T. Asuroglu, "A comprehensive evaluation of oversampling techniques for enhancing text classification performance," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-05791-7.

[15] V. B. Kobayashi, S. T. Mol, H. A. Berkers, G. Kismihók, and D. N. Den Hartog, "Text Classification for Organizational Researchers: A Tutorial," *Organ. Res. Methods*, vol. 21, no. 3, pp. 766–799, Jul. 2018, doi: 10.1177/1094428117719322.

[16] S. M. Khabour, D. Mustafa, and Q. AlThebyan, "Arabic Sentiment Analysis of Mobile Banking Services Reviews," in *2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, IEEE, Nov. 2023, pp. 1–8. doi: 10.1109/SNAMS60348.2023.10375427.

[17] A. Gasparetto, M. Marcuzzo, A. Zangari, and A. Albarelli, "A Survey on Text Classification Algorithms: From Text to Predictions," *Information (Switzerland)*, vol. 13, no. 2, Feb. 2022, doi: 10.3390/info13020083.

[18] D. Suleiman, A. Odeh, and R. Al-Sayyed, "Arabic Sentiment Analysis Using Naïve Bayes and CNN-LSTM," *Informatica (Slovenia)*, vol. 46, no. 6, pp. 79–86, 2022, doi: 10.31449/inf.v46i6.4199.

[19] C. Villavicencio, J. J. Macrohon, X. A. Inbaraj, J. H. Jeng, and J. G. Hsieh, "Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes," *Information (Switzerland)*, vol. 12, no. 5, May 2021, doi: 10.3390/info12050204.

[20] S. Ouf, K. T. Mahmoud, and M. A. Abdel-Fattah, "A proposed hybrid framework to improve the accuracy of customer churn prediction in telecom industry," *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/s40537-024-00922-9.

[21] R. Nashir, E. B. Setiawan, and D. Adytia, "The Influence of Sentiment on the Movement of Bank Mandiri (BMRI) Stock Price with Word2Vec Feature Expansion and the Naïve Bayes-Support Vector Machine (NBSVM) Classifier," in *2022 International Conference on Data Science and Its Applications, ICoDSA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 76–81. doi: 10.1109/ICoDSA55874.2022.9862919.

[22] J. Zhou, X. Yao, X. Miao, F. He, and Y. Miao, "Impact of Social Sentiment Changes on People's Consumption Behavior during COVID-19 Outbreak," in *2023 8th International Conference on Cloud Computing and Big Data Analytics, ICCCBDA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 113–119. doi: 10.1109/ICCCBDA56900.2023.10154695.

[23] J. Jeyasudha and D. Usha, "Topic Modeling and Sentimental Analysis of Tweets on Covid19 to Find the Weightage of the Popular Hashtag," 2021.

[24] PT. Bank Tabungan Negara (Persero) Tbk, "Analyst Meeting per 31 Maret-2023-unaudited," Jakarta, Apr. 2023. Accessed: May 30, 2024. [Online]. Available: https://btn.co.id/sitecore/content/BTN/tentang-btn/hubungan-investor/presentasi-perusahaan

[25] PT. Bank Tabungan Negara (Persero) Tbk, "Analyst Meeting per 30 Juni-2023-unaudited," 2023. Accessed: May 30, 2024. [Online]. Available: https://btn.co.id/sitecore/content/BTN/tentang-btn/hubungan-investor/presentasi-perusahaan

p-ISSN: 2541-1179
e-ISSN: 2581-1711
JURNAL INSTEK
INFORMATIKA SAINS DAN TEKNOLOGI

[26] PT. Bank Tabungan Negara (Persero) Tbk, "Analyst Meeting per 30 September-2023-unaudited," Jakarta, Nov. 2023. Accessed: May 30, 2024. [Online]. Available: https://btn.co.id/sitecore/content/BTN/tentang-btn/hubungan-investor/presentasi-perusahaan

[27] PT. Bank Tabungan Negara (Persero) Tbk, "Analyst Meeting per 31 Desember-2023-audited," Jakarta, Feb. 2024. Accessed: May 30, 2024. [Online]. Available: https://btn.co.id/sitecore/content/BTN/tentang-btn/hubungan-investor/presentasi-perusahaan

[28] PT. Bank Tabungan Negara (Persero) Tbk, "Laporan Tahunan 2023 PT. Bank Tabungan Negara (Persero) Tbk," Feb. 2024. Accessed: May 30, 2024. [Online]. Available: https://www.btn.co.id/id/About/Investor-Relation/Company-Report/Annual-Report