P-ISSN: 2656-1743 | E-ISSN: 2656-1735 **JURNAL RISET INFORMATIKA**
DOI: https://doi.org/10.34288/jri.v8i2.500 Vol. 8, No. 2. Maret 2026
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

SENTIMENT CLASSIFICATION OF PUBLIC PERCEPTIONS ON RP200 TRILLION HIMBARA STIMULUS USING NAÏVE BAYES

Wan Sobri Amin[-1], Muhammad Fikry[-2], Rahmad Abdillah[-3], Surya Agustian[-4]

[1234]Teknik Informatika / Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

12050116061@students.uin-suska.ac.id[-1] , muhammad.fikry@uin-suska.ac.id[-2], rahmad.abdillah@uin-suska.ac.id[-3], surya.agustian@uin-suska.ac.id[-4]

Abstract

The government's policy in the form of a fund stimulus of Rp200 trillion to the *Himpunan Bank Milik Negara* (HIMBARA) is a strategic step to maintain national economic stability and encourage real sector recovery. However, the implementation of public policy is inseparable from the response and public perception that develops on social media. This study aims to classify public sentiment towards the Rp200 trillion fund stimulus policy to Bank HIMBARA based on Instagram user comments and test the performance of the Naïve Bayes Classifier method in analyzing public policy sentiment. This study uses a quantitative approach with text mining and machine learning methods. Data in the form of 1.309 Instagram comments was collected through web scraping techniques from several online media accounts, then processed through text preprocessing and manual labeling stages into positive, neutral, and negative sentiments. Feature weighting was carried out using TF-IDF, then the data were classified using Multinomial Naïve Bayes and Complement Naïve Bayes. The results show that the Complement Naïve Bayes model achieved the best performance with an accuracy of 84%, an F1-score of 81%, and a high ROC-AUC value. These findings indicate that the majority of public sentiment toward the stimulus policy tends to be positive, and that the Naïve Bayes method is effective for social media–based sentiment analysis.

Keywords: Sentiment Analysis; Naïve Bayes Classifier; TF-IDF; Government Policy; Instagram

Abstrak

Kebijakan pemerintah berupa stimulus dana sebesar Rp200 triliun kepada Bank Himpunan Bank Milik Negara (HIMBARA) merupakan langkah strategis untuk menjaga stabilitas ekonomi nasional dan mendorong pemulihan sektor riil. Namun, implementasi kebijakan publik tidak terlepas dari respons dan persepsi masyarakat yang berkembang di media sosial. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap kebijakan stimulus dana Rp200 triliun ke Bank HIMBARA berdasarkan komentar pengguna Instagram serta menguji kinerja metode Naïve Bayes Classifier dalam analisis sentimen kebijakan publik. Penelitian ini menggunakan pendekatan kuantitatif dengan metode text mining dan machine learning. Data berupa 1.309 komentar Instagram dikumpulkan melalui teknik web scraping dari beberapa akun media daring, kemudian diproses melalui tahapan text preprocessing dan pelabelan manual menjadi sentimen positif, netral, dan negatif. Pembobotan fitur dilakukan menggunakan TF-IDF, selanjutnya data diklasifikasikan menggunakan Multinomial Naïve Bayes dan Complement Naïve Bayes. Hasil penelitian menunjukkan bahwa model Complement Naïve Bayes menghasilkan performa terbaik dengan akurasi sebesar 84%, nilai F1-score 81%, serta nilai ROC-AUC yang tinggi. Temuan ini menunjukkan bahwa mayoritas sentimen masyarakat terhadap kebijakan stimulus tersebut cenderung positif, serta metode Naïve Bayes efektif digunakan dalam analisis sentimen berbasis media sosial.

Kata kunci: *Analisis Sentimen; Naïve Bayes Classifier; TF-IDF; Kebijakan Pemerintah; Instagram*

INTRODUCTION

The fiscal stimulus policy is one of the government's strategic instruments in maintaining national economic stability, especially in conditions of global and domestic economic uncertainty (Kemenkeu RI 2025). Fiscal stimulus is directed to strengthen state spending, increase financial sector liquidity, and maintain people's purchasing power so that economic activities continue to run. Through this policy, the government seeks to minimize the impact of the economic slowdown by ensuring the availability of adequate financing for the banking sector and the real sector (Masrufah, 2022). In this context, fiscal policy not only serves as a tool for economic stabilization, but also as a mechanism for government intervention to maintain public confidence in the national financial system (Kemenkeu RI 2025).

HIMBARA, which consists of Bank Mandiri, BRI, BNI, and BTN, has a central role in the Indonesian financial system as the government's main partner in the implementation of fiscal and monetary policies (Saputri, 2023). Based on the theory of banking intermediation, banks function as institutions that collect funds and redistribute them to the public in the form of credit to encourage economic activity (Silitonga & Manda, 2022). By strengthening liquidity through the placement of stimulus funds, banks are expected to be able to expand access to financing, especially for the *usaha mikro, kecil, dan menengah* (UMKM), thus encouraging increased productivity and economic stability of the community (Konstantakopoulou, 2023). Therefore, the distribution of stimulus through HIMBARA is considered a measurable and strategic policy step in supporting national economic recovery.

On the other hand, the implementation of public policy is inseparable from public response and perception. Public perception is an important indicator to assess the level of acceptance, legitimacy, and effectiveness of a policy (Konstantakopoulou, 2023). In recent years, various government policies have often given rise to negative responses in public spaces, especially on social media, because they are considered less in favor of the needs of the community (Masrufah, 2022). However, the inauguration of Purbaya Yudhi Sadewa as Minister of Finance shows a change in the direction of fiscal policy that is perceived to be more pro-people, one of which is through the Rp200 trillion fund stimulus policy to Bank HIMBARA.

Based on the dynamics of public opinion on Instagram social media, this policy has received a relatively more positive response than previous policies. Many public comments assessed that the policy has the potential to help stabilize the economy, increase access to financing, and support sustainability of UMKM. However, the perceptions that emerge in the popular comment column do not necessarily represent the overall public sentiment. This raises the need to conduct a more objective and data-based study to comprehensively measure the tendency of public sentiment.

In this context, data-driven sentiment analysis is a relevant approach to systematically classify public opinion (Setiawan & Fathonah, 2025). The use of the Naïve Bayes Classifier method is widely used because this algorithm is known to be simple, efficient, and effective in classifying large amounts of text data (Nurainun et al., 2023). Metode Naïve Bayes bekerja berdasarkan teori probabilitas dengan asumsi independensi antar fitur, sehingga sesuai untuk menganalisis komentar publik di media sosial yang bersifat tidak terstruktur dan beragam (Naraswati et al., 2021).

Various previous studies have proven the effectiveness of the Naïve Bayes Classifier method in the analysis of public policy sentiment. Research by (Alfaridzy et al., 2025) showed that the use of Naïve Bayes with TF-IDF weighting on Instagram comments related to government budget efficiency policies was able to achieve an accuracy level of 90.74%. The findings show that this method has a reliable performance in classifying public sentiment on government policy issues (Rieuwpassa et al., 2024). And the research conducted (Putra & Putra, 2025) The resulting study with the implementation of the Naïve Bayes method showed an accuracy value of 81% indicating that the method has been successful in analyzing user sentiment. The findings in the study show that positive sentiment is more dominant and although negative sentiment is small, evaluation and improvement based on user complaints is still needed for application development.

Furthermore, the research carried out by (Hidayat, 2024) which states that sentiment analysis can be an effective tool to understand public opinions, especially netizens on the Twitter platform, related to perceptions of Puan Maharani. The highest accuracy value of the unbalanced dataset was obtained at 88.89% at the ratio of training data and test data sharing of 90:10 and the highest accuracy of the balanced dataset of 81.0% at the ratio of data distribution of 90:10.

In contrast to previous studies that applied only a single variant of Naïve Bayes to Twitter data without evaluating the contribution of each preprocessing stage, this study introduces novelty through a systematic comparison between Multinomial and Complement Naïve Bayes on a fiscal policy dataset, as well as through an ablation study to identify the impact of each preprocessing step on model performance. In addition, this research conducts a comprehensive N-gram feature comparison to analyze the effectiveness of various word-based n-gram configurations in representing sentiment patterns in social media text. Furthermore, this study specifically examines public responses to the Rp200 trillion fiscal stimulus policy, which has not yet been addressed in the sentiment classification literature.

And the research conducted Based on this description, a study is needed that examines public sentiment towards the Rp200 trillion fund stimulus policy to Bank HIMBARA in an objective and data-based manner. Therefore, this study aims to classify public sentiment based on comments on Instagram social media using the Naïve Bayes Classifier method. The results of this study are expected to provide a more accurate picture of the level of public acceptance of the stimulus policy, as well as an evaluation material for the government in formulating communication strategies and implementing economic policies in the future.

RESEARCH METHODS

Types of research

This research is an applied quantitative research with an experimental approach based on text mining and machine learning. The study aims to classify public sentiment towards the government's policy of stimulating funds of Rp200 trillion to Bank HIMBARA using the Naïve Bayes Classifier method.

Time and Place of Research

This research was conducted in 2025. Data collection was carried out online through Instagram social media from September 23 to October 27, 2025. The data processing and analysis were conducted in the researcher's computing environment.

Research Target / Subject

The objective of this study is to develop and evaluate a sentiment classification model capable of categorizing public opinion into positive, neutral, and negative classes regarding the Rp200 trillion fiscal stimulus policy to Bank HIMBARA. In addition to identifying public sentiment trends, this research also aims to conduct a systematic comparison between Multinomial and Complement Naïve Bayes, analyze the effectiveness of various n-gram feature configurations, and examine the contribution of each preprocessing stage through an ablation study approach. The research subject consists of public comments on Instagram posts discussing the policy, while the object of the research is the textual content of the comments representing public opinion toward the analyzed policy.

Procedure

The research procedure includes the stages of comment data collection, text preprocessing (cleaning, tokenization, stopword removal, and stemming), feature weighting using TF-IDF, classification process with the Naïve Bayes Classifier algorithm, and model performance evaluation.

Data, Instruments, and Data Collection Techniques

The data used in this study is secondary data in the form of public comment texts obtained from Instagram social media. The comments.

Comes from an upload that discusses government policies related to fund stimulus of IDR 200 trillion to Bank HIMBARA. This text data represents public opinion that is spontaneous and open to the policies being studied.

Data collection is carried out using the web scraping technique, which is the process of automatically retrieving comment data from the Instagram platform based on keywords and uploads relevant to the research topic (Balan Pratama et al., 2021). The data that has been collected is then selected to ensure fit for the context of the policy being analyzed and to avoid duplication and irrelevant content.

The research instruments used are software and text data processing tools, including programming languages and supporting libraries used for the data collection process, text preprocessing, and sentiment analysis. The preprocessing stage includes text cleaning, tokenization, stopword removal, and stemming to produce data that is ready for analysis (Nurrochmah et al., 2025).

The data collection technique in this study was carried out in a non-interactive manner, without directly involving respondents, so as not to

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v8i2.500
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

affect the content of the opinions analyzed. This approach allows researchers to obtain data that reflects people's perception of the government's natural perception of government policies that are the object of research.

The data used in this study consist of publicly accessible comments obtained from verified Instagram media accounts recognized by the Press Council, discussing the Rp200 trillion fiscal stimulus policy. Data collection complied with Instagram's platform policies and excluded private or restricted content. No personally identifiable information was collected, and all usernames were anonymized during data processing. The analysis focused solely on textual content relevant to the research objectives.

Data analysis technique

Data analysis was conducted using the Naïve Bayes Classifier with TF-IDF weighting. The model was implemented in Python using the Scikit-learn library and employed two variants of Naïve Bayes, namely Multinomial Naïve Bayes as the baseline model and Complement Naïve Bayes as the optimized model. Both models applied Laplace smoothing with an alpha value of 1.0 and used default parameter configurations. Feature extraction was performed using TF-IDF with customized settings. Multinomial Naïve Bayes used word-based n-gram configurations (1,1), (1,2), and (2,2) with min_df = 2, max_df = 0.95, and sublinear_tf = True. Meanwhile, Complement Naïve Bayes employed character-based n-gram features with a range of (3–6) using analyzer = "char_wb" under the same TF-IDF configuration. Model evaluation was conducted using Stratified 5-Fold Cross-Validation and a 90:10 hold-out validation scheme with random_state = 42. Model performance was evaluated using Accuracy, F1-score, and Receiver Operating Characteristic and metrik Area Under the Curve (ROC-AUC).

Language Characteristics Analysis

Researchers found various linguistic anomalies in people's comments, such as the use of non-standard words that are not listed in KBBI, slang, abbreviations, spelling variations, repetition of letters as an emphasis on emotions, and the use of emoticons and symbols that represent certain sentiments.

RESULTS AND DISCUSSION

Data collection

Data collection in this study was carried out by collecting public comments on the Instagram social media platform which discussed government policies related to the fund stimulus of Rp 200 trillion to the HIMBARA bank. The comment data was obtained from the uploads of social media accounts, namely @kumparancom, @suaradotcom, @republikaonline, and @detikcom, this social media has been verified by the press council so that the information presented has a more guaranteed level of credibility, accuracy, and accountability.

Manual Labeling

Sentiment labeling is done manually with three classes of sentiment, namely positive, neutral, and negative. The manual labeling was validated by an Indonesian teacher at SMAS Islam Raudhatul Jannah, Payakumbuh. There were 1309 comments consisting of, 704 positive, 309 neutral and 296 negative, the results of manual labeling can be seen in table 1

Table 1. Manual Labeling

User ID	Comment	Sentiment
U1	DPR mau berulah lagi ini😎	Negatif
U2	Semoga bpknya sehat ,, biasanya klo orang di pemerintahan kerjanya lurus nnti bakal di ganggu sama tikus tikus lucknat🔥 ID	Positif
U3	karakter asli nya laki^2 keliatan kalau punya punya uang	Netral
U4	Jangan masuk angin .	Netral
U5	pak tolong benahin perekonomian indonesia...demi rakyat yang selalu hidup sengsara ...	Positif
U6	Semoga bisa membawa negara ini yg lebih baik⚽😳	Positif
U7	DPR panik🔥🔥🔥	Negatif

Although the dataset shows a moderately imbalanced class distribution, this study did not apply additional resampling techniques. This decision was based on the consideration that the imbalance ratio was not extreme and on the use of Complement Naïve Bayes, which is specifically designed to handle text classification problems with imbalanced class distributions. Furthermore, model evaluation was not solely based on accuracy but also employed the Macro F1-score, which is more appropriate for imbalanced datasets as it assigns equal weight to each class.

During the manual labeling process, researchers found a number of linguistic characteristics that were not fully accommodated by the standard preprocessing stages. These characteristics include the use of non-standard words that are not listed in the KBBI, the repetition of letters as an emphasis on emotions, the use of emojis according to social media culture, and the use of punctuation marks and question sentences that affect the interpretation of sentiment. Based on these findings, adjustments were made to preprocessing rules and domain-knowledge-based features to accommodate data-specific characteristics.

Preprocessing

The text preprocessing stage is a crucial part of this study, considering that the data used comes from social media with a high level of noise (Nugraha & Siregar, 2021). Setiap tahap preprocessing dirancang untuk saling melengkapi dan terintegrasi secara sistematis (Nurwanda et al., 2024).

The cleaning and case folding process ensures that the text is free of irrelevant elements and has a uniform writing form (Naraswati et al., 2021). Normalization serves to align the variations of non-standard words and informal language of social media into a more consistent form, thereby reducing the diversity of unnecessary features (Prayugah et al., 2024). Tokenizing converts text into a computatively processable unit of word, while stopword removal preserves meaningful words by eliminating common words that don't contribute to sentiment (Alshehri & Algarni, 2023). The stemming stage then unites the suffix variations into the basic form, so that the resulting features become more concise and representative. Overall, this preprocessing feature forms a cleaner, consistent, and more relevant representation of text for use at the TF-IDF weighting and sentiment classification stages. The following stages of data processing can be seen in table 2.

Table 2. Data Preprocessing Stages

Process	Results
Raw Data	Semoga bpknya sehat ,, biasanya klo orang di pemerintahan kerjanya lurus nnti bakal di ganggu sama tikus tikus lucknat 🔥 ID
Cleaning	Semoga bpknya sehat biasanya klo orang di pemerintahan kerjanya lurus nnti bakal di ganggu sama tikus tikus lucknat 🔥 ID
Case Folding	semoga bpknya sehat biasanya klo orang di pemerintahan kerjanya lurus nnti bakal di ganggu sama tikus tikus lucknat 🔥 ID
Normalization	semoga bpknya sehat biasanya klo orang di pemerintahan kerjanya lurus nnti bakal di ganggu sama tikus tikus laknat emojisemangat
Tokenizing	['semoga', 'bpknya', 'sehat', 'biasanya', 'klo', 'orang', 'di', 'pemerintahan', 'kerjanya', 'lurus', 'nnti', 'bakal', 'di', 'ganggu', 'sama', 'tikus', 'tikus', 'lucknat', 'emojisemangat']
Stopword Removal	semoga bpknya sehat klo orang pemerintahan kerjanya lurus nnti ganggu tikus tikus laknat emojisemangat hintpositif
Stemming	moga bpknya sehat klo orang perintah kerja lurus nnti ganggu tikus tikus laknat emojisemangat hintpositif

The text preprocessing stage produces a set of text data which is then analyzed based on the word frequency distribution. The words with the highest occurrence rates were then visualized in wordcloud form for positive and negative sentiment categories, with the aim of identifying dominant topics as well as sentiment trends that appeared in the text data.

Wordcloud

Wordcloud visualization before the text preprocessing stage shows that the dominant words are still dominated by raw forms, spelling variations, repetitive words, and common tokens that do not clearly represent the meaning of sentiment, so the main topics in the data are not yet

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v8i2.500
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

specifically visible. After preprocessing, wordcloud shows significant changes, where the words appear to become more structured and informative, such as banks, people, money, credit, and economy. This shows that the preprocessing stage succeeds in reducing noise, uniting word variations, and improving the quality of text representation so that the topics discussed in the data become clearer and more relevant. In the context of text classification, high-frequency words influence the TF-IDF weighting process because they contribute to feature formation within the vector space representation. Although TF-IDF reduces the dominance of overly common terms through inverse document frequency weighting, words that consistently appear within specific sentiment classes can still provide strong discriminatory signals, the image of the WordCloud result can be seen in Figure 1.

Figure 1. Wordcloud

TF-IDF Weighting

The Term Frequency - Inverse Document Frequency (TF-IDF) method is used to convert the preprocessed text into a numeric vector (Azzahra & Mailoa, 2025). TF-IDF is able to balance the frequency of word occurrences with their importance throughout the document, so that words that appear frequently but are less informative do not dominate the classification process (Alfawas et al., 2024).

Naïve Bayes Classifier

This study applied two types of Naïve Bayes Classifier as part of a comparative analysis. The first model uses Multinomial Naïve Bayes as a general baseline model and the second model uses Complement Naïve Bayes as a comparator model to assess the extent to which feature optimization and Naïve Bayes type selection affect classification performance (Mola et al., 2025).

Multinomial Naïve Bayes is combined with n-gram-based TF-IDF weighting, which includes unigram, unigram+bigram and bigram with a 90:10 test and train data scheme. Complement Naïve Bayes is designed to improve classification performance on datasets with unbalanced class distributions. Combined with TF-IDF weighting, n-gram characters with a range of 3 - 6 characters were chosen because it was enough to represent meaningful pieces of words without causing excessive sparsity. This classification uses a 90:10 test and training data scheme.

N-gram Feature Comparison

The n-gram configuration in Multinomial Naïve Bayes consists of unigrams, unigrams+bigrams, and bigrams, each of which has different characteristics in representing the context of the text. Unigram (1,1) forms a feature of a single word so that it is able to capture the basic meaning of each word directly, unigram + bigram (1,2) combines single words and sequential word pairs to expand the context, Meanwhile, the bigram configuration alone (2,2) only utilizes word pairs as features, which in short comment data tend to appear rarely and lose the context of the meaning when standing alone, resulting in the lowest classification performance.

The comparison stage of the n-gram configuration in this study was carried out using Stratified 5 Fold Cross-Validation to ensure that the evaluation of model performance is objective and stable. In this scheme, the dataset is divided into five parts (folds) with the proportion of sentiment classes that remain balanced on each fold. The results of the comparison of the n-gram feature can be seen in Table 3.

Table 3. N-gram features Comparison

Scenario	n_features	Accuracy	F1-score
0	1.1	0,7724	0.7170
1	1.2	0,7674	0.7101
2	2.2	0,6044	0.4416

In a comparison of n-gram configurations using TF-IDF and Multinomial Naïve Bayes, unigram (1.1) produced the best performance with an accuracy value of 0.7724 and an F1-score of 0.7170. Compared to unigram+bigram (1.2) and bigram (2.2) configurations, which tend to result in a lower number of features

The use of unigrams allows the model to capture the meaning of the single word that most

often appears in public opinion, so that the representation of the text becomes simpler and more stable in describing the sentimental tendencies of the community (Kusuma, 2024).

Ablation Study

This study was conducted to measure the contribution of each preprocessing stage to the performance of the sentiment classification model and then evaluate the performance of the model using a 90:10 hold-out scheme. The stages are carried out by activating and deactivating (1 = used, 0 = not used) allowing the researcher to identify the preprocessing stages that have the most influence on the Accuracy value and F1-score, The results of the analysis can be seen in Table 4 and Table 5.

Table 4. Ablation Study Model 1

MultinomialNB									
No	Preprocessing				Cross Val	Training		Testing	
	case folding	normalisasi	stopword removal	Stemming	N-gram	Acc	F1	Acc	F1
1	1	0	0	0	1.1	0,6783	0,5790	0,7099	0,6248
2	0	1	0	0	1.1	0,6800	0,5669	0,7252	0,6398
3	0	0	1	0	1.1	0,7351	0,6704	0,7557	0,7069
4	0	0	0	1	1.1	0,6842	0,5849	0,7023	0,6103
5	1	1	0	0	1.1	0,7224	0,6413	0,7023	0,6145
6	1	0	1	0	1.1	0,7351	0,6704	0,7557	0,7069
7	1	0	0	1	1.1	0,6842	0,5849	0,7023	0,6103
8	0	1	1	0	1.1	0,7148	0,6277	0,7557	0,6859
9	0	1	0	1	1.1	0,6893	0,5816	0,7405	0,6656
10	0	0	1	1	1.1	0,7385	0,6748	0,7405	0,6781
11	0	1	1	1	1.1	0,7182	0,6332	0,7710	0,7159
12	1	0	1	1	1.1	0,7385	0,6748	0,7405	0,6781
13	1	1	0	1	1.1	0,7199	0,6367	0,7328	0,6592
14	1	1	1	0	1.1	0,7725	0,7171	0,7939	0,7541
15	1	1	1	1	1.1	0,7750	0,7211	0,7939	0,7514
16	0	0	0	0	1.1	0,6783	0,5790	0,7099	0,6248
					Best Performance			0,7939	0,7541

The results of the experiment in model 1 show that the use of stopword removal provides the most significant performance improvement compared to other preprocessing stages. where stopword removal is enabled without or with a limited combination of other preprocessing, resulting in a fairly high F1-score test value. These findings indicate that the elimination of common words that do not have sentimental meanings is able to increase the signal-to-noise ratio in the representation of text features.

The best performance was obtained when In experiment 14, which combined case folding, normalization, and stopword removal without stemming, the best performance was obtained on the test data with an accuracy of 0.7939 and an F1-score of 0.7541. These results show that normalization and the removal of stopwords have a greater contribution than stemming in the context of the data used.

Table 5. Ablation Study Model 2

Complement Naïve Bayes								
No	Preprocessing				Training		Testing	
	case folding	normalisasi	stopword removal	Stemming	Acc	F1	Acc	F1
1	1	0	0	0	0,7742	0,7277	0,8473	0,8142
2	0	1	0	0	0,7708	0,7167	0,7863	0,7563
3	0	0	1	0	0,7869	0,7524	0,8092	0,7769
4	0	0	0	1	0,7097	0,6725	0,7786	0,7535
5	1	1	0	0	0,8014	0,7584	0,8015	0,7733
6	1	0	1	0	0,7869	0,7524	0,8092	0,7769
7	1	0	0	1	0,7097	0,6725	0,7786	0,7535
8	0	1	1	0	0,7818	0,7401	0,8168	0,7838
9	0	1	0	1	0,7699	0,7184	0,7939	0,7592
10	0	0	1	1	0,7470	0,7228	0,7710	0,7446
11	0	1	1	1	0,7742	0,7320	0,7939	0,7634
12	1	0	1	1	0,7470	0,7228	0,7710	0,7446
13	1	1	0	1	0,8014	0,7589	0,7939	0,7649
14	1	1	1	0	0,8209	0,7861	0,8321	0,8028
15	1	1	1	1	0,8192	0,7856	0,8321	0,8061
16	0	0	0	0	0,7732	0,7217	0,8470	0,8122
				Best Performance			0,8473	0,8142

The results of the experiment on Model 2 show that stopword removal is the most influential component in improving performance, as seen from the increase in F1-score testing values when activating the process, either singly or in limited combinations.

However, the best performance was obtained in experiment 1, which was by only activating case folding without normalization, stopword removal, and stemming, which resulted in a test accuracy of 0.8473 and an F1-score of 0.8142.

This phenomenon can be explained by the Complement Naïve Bayes characteristic, which is designed to handle text data with an unbalanced distribution of classes by utilizing feature distribution information on a complement class. In these conditions, the removal of features through stopword removal or stemming has the potential to eliminate discriminatory information that is implicitly used by Complement Naïve Bayes to distinguish sentiment classes, so that the model actually shows optimal performance when text features are maintained more intact.

Model Comparison

The model comparison stage aims to determine the best sentiment classification model by comparing two variants of the Naïve Bayes algorithm, namely Multinomial Naïve Bayes as the baseline model and Complement Naïve Bayes as the optimization model. Comparison is carried out after all stages of preprocessing and ablation study are completed, so that the model being compared is the best configuration of each approach. The evaluation of model performance was carried out using a 90:10 hold-out validation scheme, with the main metrics in the form of accuracy and F1-score to ensure balanced performance between sentiment classes, the results of the comparison can be seen in table 6.

Table 6. Model Comparison

	Model	**Accuracy**	**F1_score**	**ROC-AUC**
0	Model 2 Complement NB	0.8473	0.8142	0.9338

1	Model 1 Multinomial NB	0.7939	0.7541	0,9301

Quantitatively, Complement Naïve Bayes with TF-IDF characters (3 - 6) achieved the best performance with an accuracy of 84.73% and an F1-score of 81.42%, while Multinomial Naïve Bayes as a baseline model produced an accuracy of 79.39% and an F1-score of 75.41%. This difference shows that Complement Naïve Bayes not only excels in overall prediction accuracy, but is also better able to maintain a balance of performance between sentiment classes. Thus, Complement Naïve Bayes was established as the best model in this study and used as the basis for the final conclusion.

In the context of public policy sentiment classification, ROC-AUC is used to evaluate the model's ability to distinguish between different sentiment categories. The results indicate that the Complement Naïve Bayes model demonstrates strong discrimination capability, with AUC values of 0.9574 for negative sentiment, 0.8904 for neutral sentiment, and 0.9537 for positive sentiment, resulting in an overall ROC-AUC of 0.9338. These findings indicate that the model effectively differentiates sentiment categories and further confirm that Complement Naïve Bayes is the most optimal model for analyzing public opinion toward government policies. The ROC curve visualization is presented in Figure 2.

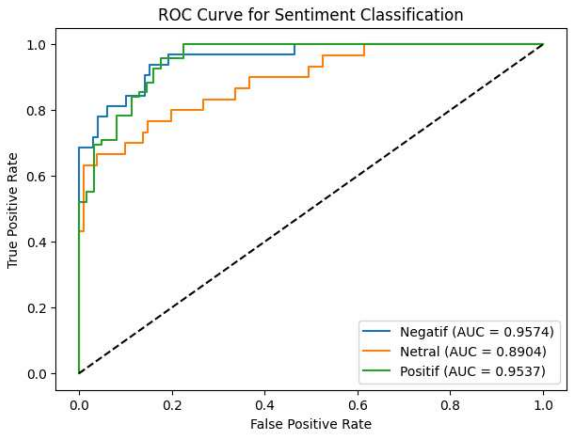

Figure 2. ROC Curve

CONCLUSIONS AND SUGGESTIONS

Conclusion

Based on the results of the study on the classification of public sentiment towards the Rp200 trillion fund stimulus policy to Bank HIMBARA using the Naïve Bayes Classifier method with TF-IDF weighting, this study succeeded in providing a more objective sentiment analysis of 1,309 Instagram comments, where positive sentiment dominated as many as 704 comments, followed by neutral 309 comments, and negative 296 comments. The comparison results showed that Complement Naïve Bayes was superior to Multinomial Naïve Bayes, with the best configuration using TF-IDF based on n-gram characters in the range of 3 - 6 in a 90:10 data sharing scheme, which achieved an accuracy of 84.73% and an F1-score of 81.42%. In addition, through ablation studies, it was found that minimal preprocessing in the form of cleaning and case folding produced optimal performance in Complement Naïve Bayes, which indicates that the complexity of preprocessing is not always directly proportional to the improvement in model performance. These findings are expected to be the basis for evaluation for the government in understanding public responses and supporting the formulation of more effective economic policy communication strategies, as well as a reference for future researchers in developing more optimal sentiment classification methods, both through model exploration, feature representation, and preprocessing strategies tailored to the characteristics of text data.

Suggestion

Although the proposed Naïve Bayes-based sentiment classification model shows good performance, the study still has some limitations. First, this study has not explicitly modeled paralinguistic features in comments, such as the use of full capital letters, the repetition of exclamation marks, and the emphasis of writing styles that can represent certain sentiments. Second, this study does not yet have a specific mechanism for dealing with sarcasm/irony, even though sarcasm often appears in public policy comments and can lead to mismatches between literal meaning and actual sentiment. As a result, some sarcastic comments have the potential to be misclassified. This limitation is an opportunity for development in further research through the exploration of paralinguistic features as well as more specific sarcasm detection approaches.

REFERENCES

Alfaridzy, M. A., Haerani, E., Jasril, & Oktavia, L. (2025). Klasifikasi Sentimen Masyarakat

Terhadap Efisiensi Anggaran Pemerintah Menggunakan Metode Naïve Bayes Classifier. *JUTECH: Journal Education and Technology*,6(1), 169–182.

Alfawas, T. I., Rahim, A., & Rudiman, R. (2024). Penerapan Fitur Ekstraksi TF-IDF untuk Analisis Sentimen Ulasan Game Bus Simulator Indonesia dengan Algoritma Naive Bayes. *Innovative: Journal Of Social Science Research*, *4*(5), 3177–3193. https://doi.org/0.31004/innovative.v4i5.13975

Alshehri, A., & Algarni, A. (2023). TF-TDA: A Novel Supervised Term Weighting Scheme for Sentiment Analysis. *Electronics (Switzerland)*, *12*(7). https://doi.org/10.3390/electronics12071632

Azzahra, W. L., & Mailoa, E. (2025). Analisis Sentimen terhadap RSUD Salatiga Menggunakan SVM dan TF-IDF. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, *6*(1), 478–489. https://doi.org/10.35870/jimik.v6i1.1208

Balan Pratama, D., Sofwan, A., & Syafei, W. A. (2021). Implementasi Teknik Web Scraping Dan Fitur Data Eksternal Pada Sistem Informasi Dosen Penelitian Dan Pengabdian Dosen Fakultas Teknik Universitas Diponegoro. *Transient: Jurnal Ilmiah Teknik Elektro*, *10*(2), 292–299. https://doi.org/10.14710/transient.v10i2.292-299

Hidayat, R. (2024). *Penerapan Naïve Bayes Classifier Dalam Klasifikasi Sentimen Publik Di Twitter Terhadap Puan Maharani* [UIN Sultan Syarif Kasim Riau]. https://repository.uin-suska.ac.id/80539/1/RIZKI HIDAYAT.pdf

Konstantakopoulou, I. (2023). Financial Intermediation, Economic Growth, and Business Cycles. *Journal of Risk and Financial Management*, *16*(12), 1–9. https://doi.org/10.3390/jrfm16120514

Kusuma, A. T. A. (2024). *Perbandingan Metode Peter Norvig, Lstm, Dan Ngram Untuk Spell Correction Bahasa Indonesia*. Universitas Islam Indonesia.

Masrufah, L. (2022). Kebijakan Moneter dan Fiskal dalam Perekonomian. *KASBANA : Jurnal Hukum Ekonomi Syariah*, *2*(1), 38–55. https://doi.org/10.53948/kasbana.v2i1.37

MasrufaUtama, P. (2022). Analisis Respons Publik di Media Sosial Terhadap Proses Legislasi RUU TNI Dalam Kerangka Demokrasi Deliberatif. *Communicology Jurnal Ilmu Komunikasi*, *13*(1). https://doi.org/https://doi.org/10.21009/COMM.034.09

Mola, S. A. S., Djawa, S. N. R., & Mauko, A. Y. (2025). *Text Mining Analisis Sentimen dengan Naïve Bayes* (S. A. S. Mola (ed.)). Kaizeng Media ublishing.

Naraswati, N. P. G., Rosmilda, D. C., Desinta, D., Khairi, F., Damaiyanti, R., & Nooraeni, R. (2021). Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification. *SISTEMASI: Jurnal Sistem Informasi*, *10*(1), 222–238.

Nugraha, S. A., & Siregar, M. U. (2021). Application of The Naive Bayes Classifier Method In The Sentiment Analysis of Twitter User About The Capital City Relocation Running Title: Application of The Naive Bayes Classifier Method. *PROC. INTERNAT. CONF. SCI. ENGIN*, *4*, 171–175.

Nurainun, N., Haerani, E., Syafria, F., & Oktavia, L. (2023). Penerapan Algoritma Naïve Bayes Classifier Dalam Klasifikasi Status Gizi Balita dengan Pengujian K-Fold Cross Validation. *Journal of Computer System and Informatics (JoSYC)*, *4*(3), 578–586. https://doi.org/10.47065/josyc.v4i3.3414

Nurrochmah, D. S., Rahaningsih, N., Dana, R. D., & Rohmat, C. L. (2025). Penerapan Algoritma Naive Bayes Dalam Analisis SentimenUlasan Aplikasi Kitalulus di Google Play Store. *Jurnal Informatika Terpadu*, *6*(1), 29–37. https://doi.org/10.54914/jit.v11i1.1544

Nurwanda, Suarna, N., & Prihartono, W. (2024). Penerapan NLP (Natural Language Processing) Dalam Analisis Sentimen Pengguna Telegram di PlayStore. *Jurnal Mahasiswa Teknik Informatika*, *8*(2), 1841–1846. https://doi.org/10.36040/jati.v8i2.8469

Prayugah, M. I., Indahyanti, U., & Ariyanti, N. (2024). Analisis Sentimen Publik atas Respons Pemerintah pada Serangan Ransomware dengan Pendekatan Machine Learning dan Smote. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, *8*(2), 333. https://doi.org/10.35145/joisie.v8i2.4764

Putra, I. G. S. D., & Putra, I. N. T. A. (2025). Implementasi Metode Naïve Bayes Pada Analisis Sentimen Pengguna Aplikasi Mobile Kita Bisa. *Jurnal Informatika dan Teknik Elektro Terapan*, *13*(2), 1202–1211. https://doi.org/10.23960/jitet.v13i2.6423

Rahayu, D. W. Y., Umam, K., & Handayani, M. R. (2025). Performance of Machine Learning

Algorithms on Imbalanced Sentiment Datasets Without Balancing Techniques. *Journal of Applied Informatics and Computing*, *9*(3), 998–1005. https://doi.org/10.30871/jaic.v9i3.9584

RI, K. (2025). *Strategi Fiskal Indonesia 2025–2026: Menjaga Stabilitas Ekonomi di Tengah Ketidakpastian Global*. Kementrian Keuangan RI Direktoral Jendral Pembendaharaan. https://djpb.kemenkeu.go.id/kppn/saumlaki/id/data-publikasi/berita-terbaru/2987-strategi-fiskal-indonesia-2025–2026-menjaga-stabilitas-ekonomi-di-tengah-ketidakpastian-global.html

Rieuwpassa, J. A., Sugito, & Widiharih, T. (2024). Implementasi Metode Naive Bayes Classifier Untuk Klasifikasi Sentimen Ulasan Pengguna Aplikasi Netflix Pada Google Play. *Jurnal Gaussian*, *12*(3), 362–371. https://doi.org/10.14710/j.gauss.12.3.362-371

Saputri, B. (2023). *Pengaruh Loan To Deposit Ratio dan Return On Asset Terhadap Harga Saham (Himpunan Bank Milik Negera) Yang Terdaftar di Bursa Efek Indonesia Periode 2018-2022* [Universitas Batanghari Jambi]. http://repository.unbari.ac.id/2697/1/Besse Saputri.pdf

Setiawan, F. A. A., & Fathonah, R. N. S. (2025). Tinjauan Sistematis Literatur : Analisis Sentimen Terhadap. *JATI : Jurnal Mahasiswa Teknik Informatika*, *9*(5), 7995–8003. https://doi.org/10.36040/jati.v9i5.14963

Silitonga, R. N., & Manda, G. S. (2022). Pengaruh Risiko Kredit dan Risiko Likuiditas terhadap Kinerja Keuangan pada Bank BUMN Periode 2015-2020. *Jurnal Maksipreneur: Manajemen, Koperasi, dan Entrepreneurship*, *12*(1), 22. https://doi.org/10.30588/jmp.v12i1.948