JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 6, No. 1. December 2023 DOI: https://doi.org/10.34288/jri.v6i1.XXX
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

PHRASE DETECTION'S IMPACT ON SENTIMENT ANALYSIS OF PUBLIC OPINION AND ONLINE MEDIA TOWARD POLITICAL FIGURES

Muhammad Irsa Nurodin-[1], Yan Puspitarani-[2*)]

Informatika
Universitas Widyatama
Bandung, Indonesia
irsa.nurodin@widyatama.ac.id[1] , yan.puspitarani@widyatama.ac.id[2*)]

(*) Corresponding Author

Abstract

Public opinion of political figures and policy significantly impacts general elections. Sentiment analysis, as a method to comprehend opinion and emotion in texts, requires the step of text preprocessing to improve data quality. However, textual data often encounters irrelevant words and ambiguous language. These conditions can impact the accuracy of sentiment analysis. Given the significance of precisely interpreting public opinion toward political figures, these issues may result in biased or inaccurate sentiment analysis outcomes. Irregular punctuation or unclear language can disturb the text's intended context, compromising sentiment analysis quality. Additionally, irrelevant words can obscure the focus of the analysis, causing fundamental changes in the original text's meaning. This research focuses on the impact of a specific preprocessing technique, namely Phrase Detection with N-Gram, on sentiment analysis of political figures. By applying this method, the study aims to detail the effects of using Bigram, Trigram, and Unigram on the quality of sentiment analysis, particularly in the context of political figures on Twitter and online media articles. This research indicates that using Bigram in Phrase Detection provides more significant results than Trigram and Unigram for most political figures at Twitter, with the highest accuracy score of 88,23%. Sentiment analysis of articles in online media also indicates various results depending on the type of N-Gram. The results indicate that using N-gram phrase detection can influence the accuracy of sentiment analysis, and the resulting accuracy values are pretty high.

Keywords: Sentiment Analysis, Text Preprocessing, N-Gram, Politic, Opinion, Online Media, Twitter

Abstrak

Opini publik terhadap tokoh politik dan kebijakan memiliki dampak besar pada pemilihan umum. Analisis sentimen, sebagai metode untuk memahami opini dan emosi dalam teks, memerlukan langkah Text Preprocessing untuk meningkatkan kualitas data. Namun, data teks seringkali menghadapi masalah seperti kata-kata yang tidak relevan dan bahasa ambigu. Kondisi-kondisi ini dapat mempengaruhi akurasi analisis sentimen. Masalah-masalah ini dapat menghasilkan hasil analisis sentimen yang bias atau tidak akurat, mengingat pentingnya interpretasi yang tepat terhadap opini publik terhadap tokoh politik. Tanda baca yang tidak teratur atau bahasa yang tidak jelas dapat mengganggu konteks asli teks, mengompromikan kualitas analisis sentimen. Selain itu, kata-kata yang tidak relevan dapat mengaburkan fokus analisis, menyebabkan perubahan mendasar dalam makna teks asli. Penelitian ini berfokus pada dampak teknik preprocessing tertentu, yaitu Phrase Detection dengan N-Gram, pada analisis sentimen terhadap tokoh politik. Dengan menerapkan metode ini, penelitian ini bertujuan untuk merinci efek penggunaan Bigram, Trigram, dan Unigram pada kualitas analisis sentimen, khususnya dalam konteks tokoh politik di Twitter dan artikel media online. Hasil penelitian ini menunjukkan bahwa penggunaan Bigram dalam Phrase Detection memberikan hasil yang lebih signifikan dibandingkan dengan Trigram dan Unigram pada sebagian besar tokoh politik di Twitter, dengan nilai akurasi tertinggi mencapai 88,23%. Analisis sentimen terhadap artikel media online juga menunjukkan variasi hasil tergantung pada jenis N-Gram. Hasil penelitian ini menunjukkan bahwa penggunaan phrase detection N-Gram dapat mempengaruhi akurasi analisis sentimen, dan nilai akurasi yang dihasilkan cukup tinggi.

Kata Kunci: Analisis sentimen, Text Preprocessing, N-Gram, politik, opini, media online, Twitter

P-ISSN: 2656-1743 | E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v6i2.268
JURNAL RISET INFORMATIKA
Vol. 6, No. 2. March 2024
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

INTRODUCTION

Public opinion on political figures and policies significantly influences election outcomes in the political arena (Andriana, 2022). Therefore, sentiment analysis becomes one of the methods used to understand opinions and emotions in texts such as tweets, reviews, or articles (Administrator, 2023). However, before conducting sentiment analysis, text Preprocessing steps must be taken to enhance data quality for more accurate and reliable results (Alwasi'a, 2020).

In reality, text data used for sentiment analysis often faces issues like irrelevant words, inconsistent punctuation, non-standard or regional language use, and ambiguous words (Alfriyanto, 2020). These conditions can affect the accuracy of sentiment analysis results and complicate result interpretation. Moreover, some words may have neutral or conflicting sentiments in the context of sentiment analysis on political figures (Alfriyanto, 2020). Therefore, specific preprocessing methods are required to uphold precision in the outcomes of the sentiment analysis.

In performing sentiment analysis on political figures, several preprocessing techniques need implementation for accurate results. These techniques include Case Folding, Filtering, Tokenization, Stemming, etc. (Khairunnisa, Adiwijaya, & Said, 2021). This study will explore the impact of applying a Text Preprocessing technique, precisely the Phrase Detection technique, using the N-Gram method (Rozi, Firdausi, & Islamiyah, 2020). N-grams are used because they have advantages in the efficiency and effectiveness of text categorization (Ogada, 2016). The research aims to compare the accuracy of Text Preprocessing with and without Phrase Detection. In this context, the accuracy comparison between text preprocessing with and without phrase detection becomes crucial. This is important because the precision and accuracy of sentiment analysis results play a crucial role in shaping the public's understanding of opinions and emotions related to political figures. When faced with the complexity of language and the variability of word meanings in the political context, using techniques such as Phrase Detection can significantly contribute to overcoming ambiguity and improving the accuracy of result interpretation.

Previous research conducted by Khairunnisa et al. explored the impact of Text Preprocessing on sentiment analysis of public comments on Twitter in the context of the COVID-19 pandemic. The study found that Text Preprocessing improved the accuracy of sentiment analysis results and maintained accurate interpretation (Khairunnisa, Adiwijaya, & Said, 2021). However, this research was limited to the COVID-19 pandemic and Twitter context. Additionally, it did not address the influence of Phrase Detection on sentiment analysis but focused on the overall impact of Text Preprocessing on sentiment analysis.

Syifa Ayu Anjani and Achmad Fauzan have also conducted similar research. This study analyzes the responses of the DIY community to the implementation of PSBB (Large-Scale Social Restrictions) in Java-Bali phase II through tweet data collected from Twitter. The data was analyzed using the N-Gram method and sentiment classification using the Naïve Bayes Classifier (NBC). The research results show a classification accuracy ranging from 80.36% to 83.04%. This study provides important insights into the complaints frequently expressed by the DIY community regarding the PSBB policy and can serve as an evaluation for the DIY government (Anjani & Fauzan, 2021). However, similar to the research by Khoirunnisa et al., this study is limited to the context of COVID-19, and the media used is only the social media platform Twitter.

In addition, Bhustomy conducted research to analyze and determine whether text data preprocessing can influence or improve data mining results, understand the detailed process of sentiment analysis in machine learning, and explore Naïve Bayes classification. The dataset analyzed pertained to film reviews. The research results showed that the Naïve Bayes classifier performed well on the Baseline data, which is clean data without text processing, with an accuracy score of 0.83993. Stop Words (0.83632) ranked second, followed by Stemming (0.82364). However, this study only compared the accuracy scores among three methods: Baseline, Stopword, and Stemming (Hakim, 2021).

Another study by Diffa et al. aimed to measure the impact of preprocessing on non-standard words and typo correction on the accuracy level of sentiment analysis on Twitter. The methods used were Support Vector Machine and Levenshtein Distance normalization. The research results indicated that sentiment analysis without correction of non-standard words and typo correction yielded better accuracy, reaching 93%, compared to the support vector machine algorithm with non-standard word correction, which achieved an accuracy of 91%. However, this study was limited to the context of PPKM Nataru, and the media used was only the Twitter social media platform (Julidhiya, 2022).

Therefore, similar research involving different contexts or platforms could provide new insights into the influence of Phrase Detection on sentiment analysis.

RESEARCH METHODS

The research method employed involves a quantitative approach as calculations are conducted to observe the accuracy results of sentiment analysis using several N-gram Phrase Detection methods (Sugiyono, 2013).

Research Target / Subject

The data used for the research consists of public opinions and online media perceptions regarding political figures. Online media data were obtained from two online platforms, namely Kompas and Liputan6. Public opinion data were also collected from various comments on the Twitter social media platform.

Opinion data were obtained by gathering information from the Twitter social media platform and two online media platforms using keywords related to the names of well-known political figures in society, such as Anies Baswedan, Basuki Tjahaja Purnama (Ahok), Ganjar Pranowo, Prabowo Subianto, Puan Maharani, and Ridwan Kamil.

Twitter data included public opinions and news uploaded on Twitter mentioning the names of the political figures. Meanwhile, online media data consisted of news articles from Kompas and Liputan6 mentioning these political figures. Therefore, some data were retrieved even if the news content did not specifically discuss the mentioned political figures, leading to differences in the characteristics of data obtained from Twitter and online media.

Opinion data were collected from the Twitter social media platform within the period starting from July 1, 2023, to September 1, 2023. Meanwhile, data from media sources were collected from April 1, 2023, to September 1, 2023. The research methodology that will be used in this research is as shown in Figure 1 below:

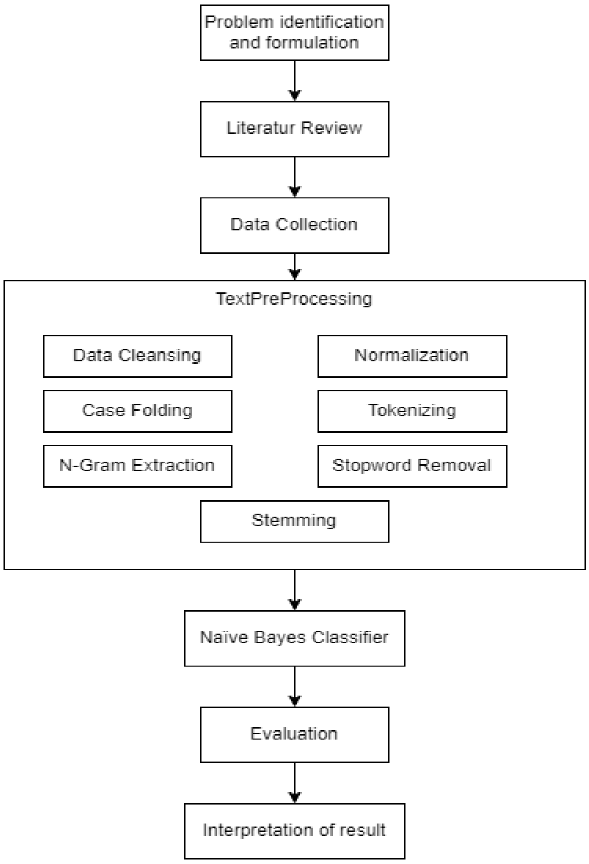

Figure 1. Research Method

Problem Identification and Formulation

In this phase, issues are identified to evaluate the public opinion trends in Indonesia towards political figures. Initial analysis results indicate an increasing monitoring of political figures in Indonesia through social media. Active users on platforms like Twitter often express their views on political figures, both positively and negatively. Based on the initial identification, the following problem formulations arise:

1. How does the influence of Text Preprocessing techniques, specifically Phrase Detection using the N-Gram method, affect the accuracy of sentiment analysis on political figures and policies?
2. What is the accuracy of comparing sentiment analysis with the application of Phrase Detection techniques and without Phrase Detection techniques on the text data?

The research's primary focus will be exploring the impact of Phrase Detection techniques (using the N-Gram method) on sentiment analysis accuracy and comparing it with situations where such techniques are not applied to text data related to political figures and policies.

P-ISSN: 2656-1743 | E-ISSN: 2656-1735 JURNAL RISET INFORMATIKA
DOI: https://doi.org/10.34288/jri.v6i2.268 Vol. 6, No. 2. March 2024
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Literature Review

The literature review is a critical phase in the research process that supports researchers in understanding the conceptual framework, assessing previous studies, and identifying areas of knowledge gaps to be explored in the planned research. In this stage, a study is conducted on various topics related to the research, such as sentiment analysis, data mining, text mining, and the Naïve Bayes Classifier.

Table 1 Literature Review 1

Journal Writter	Syifa Khairunnisa , Adiwijaya, Said Al Faraby
Journal title	*Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial* Twitter *(Studi Kasus Pandemi COVID-19)*
Page	406-414
Theory	Several challenges will be encountered when conducting sentiment analysis using Twitter data. Generally, tweet data still contains many non-standard words, such as abbreviated spellings and colloquial language. This is due to Twitter's limitation on character count, allowing a maximum of 140 characters per tweet upload. Therefore, it is necessary to perform preprocessing on tweet data as an initial step in sentiment analysis to generate better-shaped data that can be utilized in subsequent processes. Preprocessing addresses feature or attribute extraction errors and can significantly improve sentiment analysis performance.
Method	Support Vending Machine (SVM), N-Gram, Text preprocessing
Result	Based on the results of several testing scenarios conducted for sentiment analysis on COVID-19 comment tweets, it can be concluded that the best system performance is achieved when using a combination of preprocessing cleaning, stemming (without word normalization and stopword removal), and a combination of word normalization, cleaning, and stemming (without stopword removal), both resulting in the same accuracy of 77.77%.

Table 2 Literature Review 2

Journal Writter	**Syifa Ayu Anjani, Achmad Fauzan**
Journal title	*Implementasi* n-Gram *dalam Analisis Sentimen Masyarakat DIY terhadap PSBB Jawa-Bali Jilid II Menggunakan* Naive Bayes Classifier
Page	73-83
Theory	The classification process determines which reviews are positive, neutral, and hostile. Subsequently, the data is divided into training and testing data with an 80:20 ratio. Machine learning utilizes the training data to learn patterns within the data to form the classifier model.
Method	Naïve Bayes Classifier, Text preprocessing, N-Gram
Result	The research conducted by Syifa Ayu Anjani and Achmad Fauzan analyzed the public response in the Yogyakarta Special Region (DIY) to the implementation of the Large-Scale Social Restrictions (PSBB) phase II in Java-Bali through data collected from Twitter. The data was analyzed using the N-Gram method, and sentiment classification was performed using the Naïve Bayes Classifier (NBC). The research results indicated a classification accuracy ranging from 80.36% to 83.04%. This study provides valuable insights into the common complaints voiced by the community in DIY regarding the PSBB policy and can serve as an evaluation for the local government.

Data Collection

The first step in this research is to collect text data from various sources such as Twitter, online news, and articles related to political figures and policies under discussion. The text data collected must be related to political figures or policies being discussed in society. Data is collected by crawling on several online and Twitter social media platforms.

Text Preprocessing

The procedures conducted in this research are similar to the standard sentiment analysis procedures but with a slight modification in the text preprocessing stage by adding the N-Gram Extraction step after Tokenizing (Putranto, Setyawati, & Wijono, 2016). After successfully collecting the text data, the text preprocessing stage will be carried out, including several techniques such as Cleansing, Normalization, Case Folding, Tokenizing, N-Gram Extraction, Stopword Removal, and Stemming (Anugerah, 2017). These preprocessing techniques are crucial for improving data quality and eliminating irrelevant words, resulting in more accurate sentiment analysis (Asgarnezhad, Shekofteh, & Boroujeni, 2017).

1. Cleansing: Removing unnecessary characters such as emojis, numbers, and punctuation (Findawati & Rosid, 2020).
2. Normalization: Correcting non-standard words (Findawati & Rosid, 2020).
3. Case Folding: Standardizing letters to lowercase (Findawati & Rosid, 2020).
4. Tokenizing: Breaking down sentence descriptions into words (Findawati & Rosid, 2020).
5. N-Gram Extraction: Breaking sentence descriptions into phrases according to the inputted value of n (Zeniarja, Salam, & Achsanu, 2020).
6. Stopword Removal: Removing meaningless words included in stopwords (Findawati & Rosid, 2020).
7. Stemming: Finding the base word by removing affixes (Findawati & Rosid, 2020).

Naïve Bayes Classification

After completing the text preprocessing stage, the next step is the Naïve Bayes classification. In this stage, a dictionary will be created to determine whether the phrases are positive, negative, or neutral. Weighting will also be applied by calculating the tf-idf values. Subsequently, classification will be performed using the Naïve Bayes classification method (Indranandita, Susanto, & Rahmat, 2008). This algorithm operates based on the principle of conditional probability, as provided by Bayes' Theorem. Bayes' Theorem calculates the probability or likelihood of an event occurring by considering the probability of another related event that has already occurred. In simpler terms, Bayes' Theorem is a method for determining probabilities when we have information about specific other probabilities (Gandhi, 2018). Teorema Bayes is expressed mathematically in the following equation :

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad \text{..........} (1)$$

Where $P(B) \neq 0$

Evaluation

Following the classification stage, an evaluation of the results will be conducted. The evaluation stage typically involves calculating accuracy, recall, precision, and F-1 score (Yunita, 2016). In this phase, a comparison will be provided between the calculation results for Twitter data and online media data.

Interpretation of Result

A Confusion Matrix will be obtained from the previous Naïve Bayes classification at this stage. Calculations for Recall, Precision, and F1-Score will be performed in this phase.

The final stage involves interpreting the results to gain new insights into public sentiment towards political figures and policies. The research findings will be interpreted and analyzed to generate knowledge. This stage will include the presentation of analysis results and conclusions and recommendations for future research development.

RESULTS AND DISCUSSION

Data Collection

Details from the Twitter and online media data scraping results can be viewed in Table 3 and Table 4.

Table 3 The details of the data from Twitter

No.	Name of Figure	Total
1	Ahok	585
2	Anies	5.000
3	Ganjar	5.000
4	Prabowo	5.000
5	Puan	1.852
6	Ridwan kamil	2.380
	Total	19.817

Table 3 lists the total data obtained through Twitter, which amounts to 19,817 data points.

Table 4 The details of the data from Online Media

No	Sumber	Tokoh	Jumlah Data
1	Kompas	Ahok	300
		Anies	300
		Ganjar	300
		Prabowo	300
		Puan	300
		Ridwan	300
		Sub Total	1800
2	Liputan6	Ahok	300
		Anies	300
		Ganjar	300
		Prabowo	300
		Puan	300
		Ridwan	300
		Sub Total	1800
		Total Data	3600

P-ISSN: 2656-1743 | E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v6i2.268

Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Table 4 presents a list of the data obtained through online media, which is a total of 3600 data.

Text Preprocessing

The obtained data is derived from scraping Twitter and online media, which then undergoes text preprocessing steps before further analysis (Putranto, Setyawati, & Wijono, 2016). The Text Preprocessing process uses Python programming, utilizing various libraries such as Natural Language Toolkit (NLTK) and Scikit-learn. In the tokenization stage, a slight modification is made by incorporating the N-Gram algorithm (Putranto, Setyawati, & Wijono, 2016). The Text Preprocessing flow is shown in Figure 2.

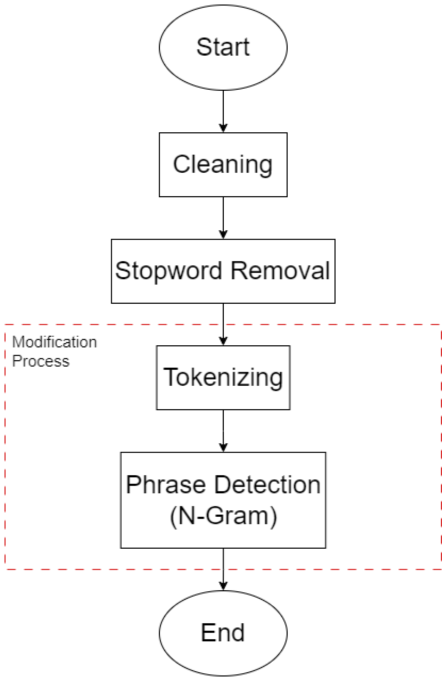

Figure 2. Text preprocessing Process

Table 5. Example of Twitter Data Text Preprocessing Results at Each Stage.

Data Grouping	Tweets atau apaalah itu kasih judul
Basic Data	*@David68159347 Ini demi bangsa dan Negara, Indonesia itu besar, butuh pemimpin yg handal dan yg sangat berpengalaman siapa dia? Ya Prabowo Subianto lah.. masa pemimpin berikutnya hanya petugas partai..* 😂
Cleaning + Case Folding	*ini demi bangsa dan negara indonesia itu besar butuh pemimpin yg handal dan yg sangat berpengalaman siapa dia ya prabowo subianto lah masa pemimpin berikutnya hanya petugas partai*
Stopword	*bangsa negara indonesia butuh pemimpin yg handal yg berpengalaman ya prabowo subianto pemimpin petugas partai*
Tokenization	*['bangsa', 'negara', 'indonesia', 'butuh', 'pemimpin', 'yg', 'handal', 'yg', 'berpengalaman', 'ya', 'prabowo', 'subianto', 'pemimpin', 'petugas', 'partai']*
N-Gram	*['bangsa negara indonesia', 'negara indonesia butuh', 'indonesia butuh pemimpin', 'butuh pemimpin yg', 'pemimpin yg handal', 'yg handal yg', 'handal yg berpengalaman', 'yg berpengalaman ya', 'berpengalaman ya prabowo', 'ya prabowo subianto', 'prabowo subianto pemimpin', 'subianto pemimpin petugas', 'pemimpin petugas partai']*
Stemming	*bangsa negara indonesia negara indonesia butuh indonesia butuh pimpin butuh pimpin yg pimpin yg handal yg handal yg handal yg alam yg alam ya alam ya prabowo ya prabowo subianto prabowo subianto pimpin subianto pimpin tugas pimpin tugas partai*

Naive Bayes Classification

Before calculating sentiment scores, the data previously retrieved through crawling is updated by adding labels related to the sentiment types of the obtained comments.

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v6i1.XXX
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Table 6 Addition of Label to Data (Sentiment Column)

source	content	date	username	sentiment
twitter	*Selain itu di sela sela kegiatannya Menhan RI Prabowo menyempatkan diri berkunjung ke Museum Keprajuritan Indonesia di Taman Mini Indonesia Indah (TMII). #UntukAnakIndonesia #PrabowoNeruskePakdhe Prabowo Subianto*	2023-07-06	s	*Positif*
Twitter	*Selain itu di sela sela kegiatannya Menhan RI Prabowo menyempatkan diri berkunjung ke Museum Keprajuritan Indonesia di Taman Mini Indonesia Indah (TMII). #UntukAnakIndonesia #PrabowoNeruskePakdhe Prabowo Subianto*	2023-07-06T08:49:09Z	alfisahdinardi	*Positif*
twitter	*@jagokan57083101 @CutSarina5 @aniesbaswedan Kecerdasan seorang Prabowo sudah sangat nyata, dari semenjak beliau jadi prajurit, hingga membesarkan partai Gerindra dan pernah mengorbitkan Ahok Jokowi, Ridwan Kamil bahkan Anis sandi. Dan karna kecerdasan dannkecintaannyalah yg membuat Prabowo Subianto seperti sekarang* 🙏	2023-07-06T17:25:35Z	sbram01	*Positif*

The labeling process is done manually for both Twitter and online media data. This is done because human decisions in the context of sentiment analysis are often more accurate, relevant, and reliable thanking specific algorithms (Catalan, 2022). The Naïve Bayes classification process flow is shown in Figure 4.

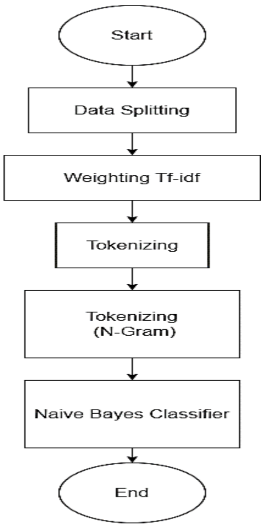

Figure 3 Naïve Bayes classification process

Evaluation

After performing Naïve Bayes classification on Twitter and online media data, the next step is calculating accuracy, precision, recall, and F1-Score values from the data. Recall is the ratio of accurate positive predictions to true positive data (Arthana, 2019). Precision is the ratio of true positive predictions compared to the overall predicted positive results (Arthana, 2019). The F1 Score is the weighted average comparison of precision and recall (Arthana, 2019). The evaluation results in process flow are shown in Figure 5.

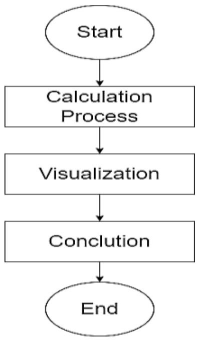

Figure 4 The evaluation results process

The provided code will generate output through a Confusion Matrix and an evaluation model diagram. The Confusion Matrix shows the number of correct and incorrect prediction values compared to the actual data, with the X-axis representing the predicted labels and the Y-axis representing the actual data labels (Kumar & Batut, 2019). The values obtained from this Confusion Matrix can be used to calculate various evaluation metrics such as accuracy, precision, recall, and F1-Score to assess the performance of the classification model (Wilianto, 2021). Here is an example of a Confusion Matrix and the resulting evaluation model.

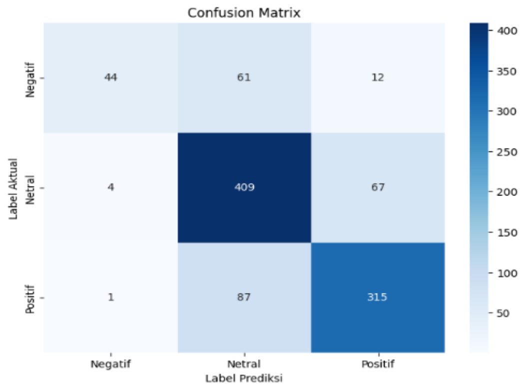

Figure 5 Confusion Matrix for analysis with Bigram method on Anies data Twitter media

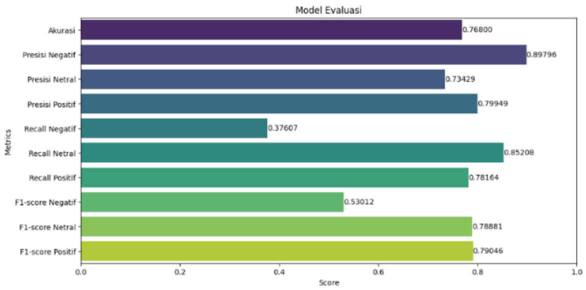

Figure 6 Evaluation Model for analysis with Bigram method on Anies data from Twitter media.

Summary of Evaluation

The following tables will present a summary of the evaluation results for the confusion matrix and obtained evaluation model.

Table 7 Comparison of Accuracy Values among N-Gram Methods for Twitter Data

Name of the Figure	Accuracy		
	N=1	N=2	N=3
Ahok	0,71795	0,72650	0,70085
Anies	0,75000	0,76800	0,76400
Ganjar	0,71900	0,74600	0,74600
Prabowo	0,86600	0,88200	0,87700
Puan	0,72507	0,73585	0,72776
Ridwan	0,88025	0,88235	0,88235

Table 8 Comparison of Accuracy Values among N-Gram Methods for Online Media Data

Name of Figure	Accuracy (Liputan6)			Accuracy (Kompas)		
	N=1	N=2	N=3	N=1	N=2	N=3
Ahok	0,62	0,58	0,58	0,77	0,70	0,70
Anies	0,80	0,77	0,77	0,88	0,78	0,75
Ganjar	0,87	0,78	0,78	0,63	0,67	0,67
Prabowo	0,72	0,75	0,75	0,82	0,80	0,82
Puan	0,87	0,87	0,87	0,80	0,83	0,82
Ridwan	0,80	0,78	0,78	0,82	0,88	0,88

Note :

: Highest Accuracy Value

Overall, the analysis of Tables 4 and 5 indicates that using the Bigram method (N-Gram with N=2) produces higher accuracy values than the Trigram and Unigram methods. This suggests that the simultaneous use of two-word combinations positively impacts the model's performance in classifying data.

After evaluating the accuracy values, the following calculation assesses the Recall, Precision, and F1-Score values.

Table 9 Comparison of average Recall, Precision, and F1-Score values based on the utilized N-Gram for Twitter data

Name of Figure	*Unigram*			*Bigram*			*Trigram*		
	Rec.	Prec.	F1	Rec.	Prec.	F1	Rec.	Prec.	F1
Ahok	0.54	0.74	0.56	0.54	0.81	0.57	0.51	0.76	0.54
Anies	0.67	0.77	0.70	0.67	0.81	0.70	0.67	0.82	0.71
Ganjar	0.56	0.70	0.60	0.57	0.77	0.61	0.55	0.77	0.59
Prabowo	0.66	0.77	0.69	0.68	0.87	0.72	0.71	0.83	0.71
Puan	0.62	0.69	0.64	0.62	0.75	0.60	0.59	0.74	0.61
Ridwan	0.54	0.69	0.56	0.48	0.52	0.49	0.49	0.85	0.51

Note :

: Highest *Recall* value

: Highest Precision value

: Highest *F1-Score* value

Table 9 provides a comparison of the average Recall, Precision, and F1-Score values based on the use of N-Grams (Unigram, Bigram, and Trigram) for six public figures, namely Ahok, Anies, Ganjar, Prabowo, Puan, and Ridwan in sentiment analysis. Overall, it shows that Bigram has better Recall, Precision, and F1-Score values than the other methods.

Table 10 Comparison of average Recall, Precision, and F1-Score values based on the utilized N-Gram for Liputan6 data

Name of Figure	*Unigram*			*Bigram*			*Trigram*		
	Rec.	Prec.	F1	Rec.	Prec.	F1	Rec.	Prec.	F1
Ahok	0.65	0.62	0.61	0.59	0.57	0.57	0.60	0.57	0.57
Anies	0.60	0.54	0.56	0.57	0.54	0.55	0.57	0.54	0.55
Ganjar	0.58	0.58	0.57	0.50	0.51	0.51	0.50	0.51	0.51
Prabowo	0.61	0.62	0.55	0.56	0.62	0.55	0.56	0.62	0.56
Puan	0.44	0.46	0.44	0.39	0.46	0.40	0.39	0.46	0.40
Ridwan	0.38	0.49	0.38	0.35	0.43	0.34	0.35	0.43	0.34

Note :

: Highest *Recall* value

: Highest Precision value

: Highest *F1-Score* value

Table 10 provides a comparison of the average Recall, Precision, and F1-Score values based on the use of N-Grams (Unigram, Bigram, and Trigram) for six public figures, namely Ahok, Anies, Ganjar, Prabowo, Puan, and Ridwan in sentiment

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v6i1.XXX
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

analysis. Overall, it shows that Unigram has better Recall, Precision, and F1-Score values than the other methods.

Table 11 Comparison of average Recall, Precision, and F1-Score values based on the utilized N-Gram for Kompas data

Name of Figure	*Unigram*			*Bigram*			*Trigram*		
	Rec.	Prec.	F1	Rec.	Prec.	F1	Rec.	Prec.	F1
Ahok	0.56	0.87	0.61	0.43	0.56	0.42	0.43	0.46	0.42
Anies	0.56	0.56	0.56	0.67	0.73	0.68	0.65	0.72	0.66
Ganjar	0.45	0.41	0.43	0.47	0.44	0.45	0.47	0.43	0.45
Prabowo	0.57	0.54	0.56	0.55	0.54	0.62	0.56	0.56	0.55
Puan	0.86	0.69	0.73	0.72	0.89	0.78	0.71	0.88	0.76
Ridwan	0.50	0.47	0.49	0.56	0.56	0.56	0.56	0.56	0.56

Note :
: Highest *Recall* value
: Highest Precision value
: Highest *F1-Score* value

Table 11 provides a comparison of the average Recall, Precision, and F1-Score values based on the use of N-Grams (Unigram, Bigram, and Trigram) for six public figures, namely Ahok, Anies, Ganjar, Prabowo, Puan, and Ridwan in sentiment analysis. Overall, it shows that Bigram has better Recall, Precision, and F1-Score values than the other methods.

CONCLUSIONS AND SUGGESTIONS

Conclusion

This research shows that the N-Gram Phrase Detection method provides varying results for analyzing public opinion and online media sentiment towards political figures, especially using the Unigram, Bigram and Trigram approaches. Some show better results using the unigram method, and some use the bigram method. Some get better results if they use the trigram method. This occurs due to several factors, including differences in the characteristics of the data used in online media and Twitter. Twitter data tends to be shorter and uses informal language, while online media data is in the form of longer paragraphs and formal language. Data per public figure used in online media is less than 500, while Twitter data per public figure exceeds 500. However, the sentiment analysis results on Twitter data show that using Bigram tends to yield more significant results.

In contrast, for online media data, the values of accuracy, recall, precision, and F1-Score vary for each number and type of N-Gram used. This shows that the choice of N-Gram type can influence the sentiment analysis results. For data from Twitter, the highest accuracy value using Bigram was obtained during sentiment analysis on the figures Ridwan Kamil and Prabowo with an accuracy of 88.235%. Meanwhile, the lowest accuracy value was obtained when analyzing Ahok's figure using Trigram, with an accuracy of 70%. For data sourced from Online Media, the highest accuracy value was obtained when analyzing the sentiment of the character Ridwan Kamil in Kompas media with a value of 88% using the Bigram and Trigram approach, while the lowest accuracy value was obtained when analyzing the sentiment of the character Ahok in Media Liputan6 with a value 58% use the Bigram and Trigram approach.

Suggestion

1. For users:
- Conduct a more in-depth data analysis and specific sentiments towards those figures.
- Understand the social, political, or policy context to enrich the understanding of sentiment analysis results.
- Ensure that the dataset reflects the diversity of opinions and sentiments from various sources to obtain more representative results.
- Emphasize research that investigates public opinions from various platforms to provide a more comprehensive overview.

2. For policymakers:
- Respond to research findings with appropriate policies if significant sentiment patterns towards political figures can influence policies.
- Collaborate with the relevant political figures to gain their perspectives on research findings and understand their impact on political policies.

3. For future researchers:
- Explore other sentiment analysis methods that can be used simultaneously or as alternatives to N-Gram phrase detection, such as Machine Learning-Based Sentiment Analysis and deep Learning with Recurrent Neural Networks (RNN). Comparing these methods can provide a deeper insight into their strengths and weaknesses.

REFERENCES

Administrator. (2023, May 15). *Cara Menggali Insight Pelanggan dengan Analisis Sentimen*. Retrieved from ivosight: https://ivosights.com/read/artikel/analisis-sentimen-cara-menggali-insight-pelanggan-dengan

Alfriyanto, M. L. (2020). *Analisis Sentimen Terhadap Operator Seluler Di Media Sosial Twitter*

Menggunakan Metode Klasifikasi Naïve Bayes Dan Metode Topsis. Gresik: Universitas Muhammadiyah Gresik.

Alwasi'a, A. (2020). *Analisis Sentimen Pada Review Aplikasi Berita Online Menggunakan Metode Maximum Entropy.* Yogyakarta: Universitas Islam Indonesia.

Andriana, N. (2022). Pandangan Partai Politik Terhadap Media Sosial. *Jurnal Penelitian Politik*, 51-65.

Anjani, S. A., & Fauzan, A. (2021). Implementasi n-Gram dalam Analisis Sentimen Masyarakat DIY Terhadap PSBB Jawa-Bali Jilid II Menggunakan Naive Bayes Classifier. *Statistika, 21*(2), 73-83.

Anugerah, F. (2017). *Perbaikan Kinerja Praproses Karakter Berulang Dalam Mengenali Kata Pada Klasifikasi Sentimen Berbahasa Indonesia.* Surabaya: Institut Teknologi Sepuluh Nopember.

Arthana, R. (2019, April 05). *Mengenal Accuracy, Precision, Recall dan Specificity serta yang diprioritaskan dalam Machine Learning*. (Medium) Retrieved May 28, 2023, from https://rey1024.medium.com/mengenal-accuracy-precission-recall-dan-specificity-serta-yang-diprioritaskan-b79ff4d77de8

Asgarnezhad, R., Shekofteh, M., & Boroujeni, F. (2017). Improving diagnosis of diabetes mellitus using combination of preprocessing techniques. *Journal of Theoretical and Applied Information Technology*, 2889–2895.

Catalan, N. (2022, December 14). *Pelabelan Data Otomatis vs Pelabelan Data Manual: Apa Bedanya?* Retrieved from tasq.ai: https://www.tasq.ai/blog/automated-data-labeling-vs-manual-data-labeling/

Findawati , Y., & Rosid, M. A. (2020). *Buku Ajar Text Mining.* Sidoarjo: UMSIDA Press.

Gandhi, R. (2018, May 5). *Naive Bayes Classifier*. (Towards Data Science) Retrieved June 04, 2023, from https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

Hakim, B. (2021). Analisa Sentimen Data Text Preprocessing Pada Data Mining Dengan Menggunakan Machine Learning. *Journal of Business and Audit Information Systems, 4*(2), 16-22.

Indranandita, A., Susanto, B., & Rahmat, A. (2008). Sistem Klasifikasi dan Pencarian Jurnal dengan Menggunakan Metode Naive Bayes dan Vector Space Model. *Jurnal Informatika, 4*, 10-18.

Julidhiya, D. A. (2022). *Pengaruh Pre-processing Terhadap Analisis Sentimen Pada Media Sosial Twitter dengan Perbaikan Kata Tidak Baku Dan Typo Correction.* Bandung: Universitas Komputer Indonesia.

Khairunnisa, S., Adiwijaya, & Said, A. (2021). Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19). *Jurnal Media Informatika Budidarma*, 406-414.

Kumar, A., & Batut, B. (2019, March 7). *Machine learning: classification and regression*. Retrieved from Galaxy Training: https://training.galaxyproject.org/training-material/topics/statistics/tutorials/classification_regression/tutorial.html

Putranto, H. A., Setyawati, O., & Wijono. (2016). Pengaruh Phrase Detection dengan POS-Tagger terhadap Akurasi Klasifikasi Sentimen menggunakan SVM. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, 252-259.

Rozi, I. F., Firdausi, A. T., & Islamiyah, K. (2020). Analisis Sentimen Pada Twitter Mengenai Pasca Bencana Menggunakan Metode Naïve Bayes Dengan Fitur N-Gram. *JIP (Jurnal Informatika Polinema)*, 33-39.

Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif Dan R&D.* Bandung: ALFABETA, CV.

Wilianto, K. (2021, October 22). *Evaluation Metrics pada Computer Vision dari Klasifikasi hingga Deteksi Objek*. Retrieved from Medium: https://medium.com/data-folks-indonesia/evaluation-metrics-pada-computer-vision-dari-klasifikasi-hingga-deteksi-objek-5049d3fd90d2

Yunita, N. (2016). Analisis Sentimen Berita Artis Dengan Menggunakan Algoritma Support Vector Machine dan Particle Swarm Optimazion. *Jurnal Sistem Informasi STMIK Antar Bangsa*, 104-112.

Zeniarja, J., Salam, A., & Achsanu, I. (2020). Sistem Koreksi Jawaban Esai Otomatis (E-Valuation) dengan Vector Space Model pada Computer Based Test (CBT). *SEMINAR NASIONAL Dinamika Informatika 2020 Universitas PGRI Yogyakarta*, 91-96.