

ANALYSIS OF PUBLIC SENTIMENT TOWARDS 2024 PRESIDENTIAL CANDIDACY USING NAÏVE BAYES ALGORITHM

Rianggi Silvi Anti¹, Nanang Ruhyana²

Sains Data
Fakultas Teknologi Informasi
Universitas Nusa Mandiri
15210018@nusamandiri.ac.id¹, nanang.ngy@nusamandiri.ac.id²

Abstract

This study analyzes public sentiment towards presidential nominations using text mining techniques and machine learning. The dataset consists of 670 tweets collected from social media. The analysis process includes a data pre-processing phase, encompassing text cleaning, case folding, tokenization, stopword removal, and stemming using the Sastrawi library for the Indonesian language. Sentiment labeling was performed using NLTK's SentimentIntensityAnalyzer, categorizing tweets into positive, negative, or neutral sentiments. The analysis results reveal the sentiment distribution among the analyzed tweets. Data modeling was performed using the Naive Bayes algorithm, which achieved an accuracy of 97.78% on the Iris dataset as an implementation example. The confusion matrix and classification report demonstrate the model's excellent performance in distinguishing sentiment classes. This research provides insights into public opinion regarding presidential nominations and demonstrates the effectiveness of text mining techniques and machine learning in sentiment analysis. The method can be applied to understand public opinion trends in other political and social contexts

Keywords: sentiment analysis, text mining, Naive Bayes, presidential nomination, social media

Abstrak

Penelitian ini menganalisis sentimen publik terhadap pencalonan presiden menggunakan teknik text mining dan pembelajaran mesin. Dataset terdiri dari 666 tweet yang diambil dari media sosial. Proses analisis meliputi tahap pre-processing data, yang mencakup pembersihan teks, case folding, tokenisasi, penghapusan stopwords, dan stemming menggunakan library Sastrawi untuk Bahasa Indonesia. Pelabelan sentimen dilakukan menggunakan NLTK's SentimentIntensityAnalyzer, mengkategorikan tweet menjadi sentimen positif dan sentimen negatif. Hasil analisis menunjukkan distribusi sentimen di antara tweet yang dianalisis. Pemodelan data dilakukan menggunakan algoritma Naive Bayes, yang mencapai akurasi 97.78% pada dataset iris sebagai contoh implementasi. Matriks konfusi dan laporan klasifikasi menunjukkan kinerja model yang sangat baik dalam membedakan kelas-kelas sentimen. Penelitian ini memberikan wawasan tentang opini publik terkait pencalonan presiden dan mendemonstrasikan efektivitas teknik text mining dan pembelajaran mesin dalam analisis sentimen. Metode ini dapat diterapkan untuk memahami tren opini publik dalam konteks politik dan sosial lainnya

Kata kunci: analisis sentimen, text mining, Naive Bayes, pencalonan presiden, media sosial

INTRODUCTION

The state is an organisation of coercive and monopolising power. The state also has the responsibility to look after the interests of the people, has a territory, and the necessary equipment to achieve the state's goals. (Permana et al., 2023). General elections to elect the head of state show that Indonesia is a democracy. The election of presidential candidates held every 5

years is part of the democratic process, especially in Indonesia. (Duei Putri et al., 2022). A politician who wants to run for president will certainly check or consider his popularity based on public opinion. Any political figure nominated to be head of state will consider their popularity based on public opinion. (Putri Nardilasari et al., 2023). Since the announcement of the 2019 Indonesian presidential candidates by each

political party, the names of the candidates have been widely discussed, especially on social media.

The 2024 presidential election in Indonesia has triggered a high intensity of public discussion, especially on social media. Sentiment analysis of these online conversations is important to understand people's perceptions and preferences towards the presidential candidates. (Al-Areef & Saputra, 2023) Today, social media has gone beyond its original role as a means of communicating with friends and family, and has become a platform for various aspects of life. Social media users not only interact socially, but also use the platform for transactions, education, and as a promotional tool, including for political parties and presidential campaigns (Fikri Haikal & Indra, 2024). Twitter is one of the largest social networking sites used by millions of users every day, where they share their thoughts and feelings on various topics. Twitter is widely used by people to express their opinions. Twitter is one example of a very popular social media network used by the younger generation as well as the older generation. On this social networking platform, a wide range of people are on the same platform, from world-famous figures to ordinary users, such as artists, sports athletes, leaders of organizations/companies, regular users (ordinary people), and certain political figures (Presidents, government councils, ministers, political party leaders, and other important political figures) (Luthfiansyah Dan et al., 2023).

This research aims to analyse public sentiment towards the 2024 presidential candidacy in Indonesia using Naïve Bayes algorithm. This algorithm was chosen due to its effective ability to classify text, especially on large datasets derived from social media such as Twitter. Naïve Bayes has been proven to be able to classify public opinion into positive, negative and neutral sentiment categories with a high degree of accuracy. For example, research by (Priansyah & Sutabri, 2024) showed that this algorithm can process real-time data from Twitter to understand the public's views on the outcome of the 2024 Indonesian General Election, making a positive contribution to the development of communication strategies and policies that are more responsive to the needs of the community. In addition, the study by (Gunawan & Purnama, n.d.) also applied the Naïve Bayes algorithm to analyse Twitter users' sentiment towards Indonesia's 2024 presidential candidates, which provided deep insights into the public's perception of the candidates in question. Thus, the use of Naïve Bayes algorithm in social media sentiment analysis is an appropriate

approach to understand public preferences and perceptions in the context of presidential elections.

The sentiment analysis process in this research includes the stages of collecting data from social media platforms, pre-processing the data, training the Naïve Bayes model, applying the model to a larger dataset, and analyzing the results obtained. In addition, this research will also observe the dynamics of changes in public sentiment over time, particularly in response to important events or controversial statements from presidential candidates. Sentiment analysis is a crucial tool in monitoring public opinion on social media, especially in the context of the 2024 presidential election, where people are actively voicing their views on the candidates.

Sentiment analysis, also known as opinion mining, is an automated process for understanding, extracting and processing textual data to obtain information about the sentiment contained in an opinion. The rapid development in sentiment analysis research and applications is fuelled by the immense influence and benefits it offers in various fields, including politics. For example, research by (Haque et al., 2022) analyzed approximately 17.5 million tweets related to the 2020 US presidential election, revealing that public sentiment can change significantly in response to certain events, as well as showing the difference in sentiment between the Twitter and YouTube platforms.

Furthermore, a study by (Yin et al., 2020) highlighted the importance of monitoring topic dynamics and sentiment on social media during the COVID-19 pandemic, finding that public sentiment can fluctuate significantly depending on the topics discussed and events occurring. This finding shows that sentiment analysis can provide deep insights into people's perceptions and emotions towards certain issues.

In the context of the 2024 presidential election in Indonesia, sentiment analysis becomes even more relevant given the high activity of political discussions on social media. As expressed by (Al-Areef & Saputra, 2023), Although the presidential election is still in the future, there is already a wide range of public sentiment, both positive and negative, towards the presidential candidates. However, for people who do not keep up with politics or are not active on social media, they may not have a clear preference for a particular presidential candidate.

Thus, the application of sentiment analysis using the Naïve Bayes algorithm in this study is expected to make a significant contribution in understanding public perceptions of 2024 presidential candidates, as well as

assisting in formulating political communication strategies that are more effective and responsive to the dynamics of public opinion.

The results of sentiment analysis are expected to provide an objective picture of public perceptions of each presidential candidate in the 2024 elections. By utilizing the Naïve Bayes algorithm, this research aims to identify factors that influence public opinion, such as current political issues, candidate performance, and campaign dynamics on social media. The findings can be a valuable source of information for campaign teams, political analysts, the media, and the general public in understanding the dynamics of public opinion ahead of presidential elections. For example, research by (Vindua & Zailani, 2023) shows that sentiment analysis can reveal changes in public perception of presidential candidates along with the development of certain political issues. In addition, a study by (Prayoga Siswono et al., 2024a) underlines the importance of sentiment analysis in evaluating public responses to presidential inaugurations, which can affect legitimacy and trust in elected leaders. As such, the integration of sentiment analysis in the electoral process can increase transparency and public participation in democracy.

While the Naïve Bayes method has the advantage of computational efficiency and reasonably good accuracy in text classification tasks, there are a number of limitations in its application to sentiment analysis. The main challenges include difficulties in recognising sarcasm, capturing cultural nuances in language, as well as avoiding biases that may arise from unbalanced training data. Manually determining sentiment polarity, whether an opinion is positive or negative, is possible, but inefficient when applied to the large-scale data generated from social media platforms every day. Therefore, the implementation of machine learning methods such as Naïve Bayes Classifier is a very appropriate approach to automate the sentiment classification process.

A number of studies have applied this method to analyze public opinion towards Indonesia's 2024 presidential candidates. For example, research by (Prayoga Siswono et al., 2024) applied the Naïve Bayes algorithm in analysing public sentiment towards the 2024 Indonesian Presidential inauguration through YouTube comments, with the results showing a significant distribution of positive, negative, and neutral sentiments.

In addition, studies by (Fauza Majbur et al., 2024) successfully built a Naïve Bayes-based

sentiment analysis system that is able to identify public perceptions of the 2024 presidential election on X social media, with an average accuracy of 82.33% through the k-fold cross-validation technique.

Research by (Yudha Patria et al., 2025) also compared the performance of Naïve Bayes Classifier with Support Vector Machine in analysing public sentiment after the 2024 presidential election, showing that Naïve Bayes gave an accuracy of 62.41%, slightly higher than SVM.

Based on previous research entitled 'Sentiment Analysis of 2024 Presidential Candidates with Multinomial Naïve Bayes Algorithm and SMOTE Oversampling', this study focuses on overcoming data class imbalance and improving classification accuracy. The results show that the application of SMOTE significantly improves classification performance: Anies dataset accuracy increased from 65.43%-66.66% to 89.28%-90.47%, while Prabowo dataset increased from 81.65%-83.53% to 89.28%-90.47%. The N-gram technique shows varying effectiveness: 1-gram is optimal for the Anies dataset, while 2-gram is more effective for Prabowo (Permana et al., 2023)

Previous research has shown that machine learning can be used to conduct sentiment analysis on social media to determine how people feel about debatable issues. Based on this and the context of the problem, research was conducted to analyse customer sentiment towards the 2024 presidential candidacy on Twitter social media. The purpose of this research is to know the implementation of sentiment analysis with the Naïve Bayes and Support Vector Machine methods and to know the performance level of the accuracy results on both algorithms.

For example, research by (Yusuf Ramadhan Nasution et al., 2024) applied the Naïve Bayes algorithm to analyse public sentiment on Twitter regarding the issue of postponing the 2024 election. The results show that this method is effective in classifying public opinion based on its polarity, with an accuracy of 91.67% using TF-IDF technique and lexicon-based approach. Previous research has shown that machine learning can be used to conduct sentiment analysis on social media to determine how people feel about debatable issues. Based on this and the context of the problem, research was conducted to analyse customer sentiment towards the 2024 presidential candidacy on Twitter social media. The purpose of this research is to know the implementation of sentiment analysis with the Naïve Bayes and

Support Vector Machine methods and to know the performance level of the accuracy results on both algorithms.

Another study by (Nugroho et al., 2024) used the Naïve Bayes algorithm to analyse public sentiment towards alleged election violations in 2024 based on Twitter data. This study successfully identified public perceptions of the issue, demonstrating the effectiveness of Naïve Bayes in handling complex opinion data.

Research by (Afuan et al., 2025) optimised sentiment analysis of the 2024 Indonesian Presidential inauguration on platform X (formerly Twitter) by combining Naïve Bayes Classifier and SMOTE technique. The results showed an increase in accuracy from 98% to 99%, highlighting the importance of handling class imbalance in sentiment data.

In addition, (Aurelly Claudia Budianto et al., 2024) compared the performance of Naïve Bayes and K-Nearest Neighbor (KNN) in analysing sentiment towards the 2024 election results on X social media. Naïve Bayes achieved 85% accuracy, while KNN achieved 68%, showing the superiority of Naïve Bayes in sentiment classification in this context.

Finally, research by (Subowo & Alzami, 2024) applied Naïve Bayes with TF-IDF technique to analyse Twitter data related to the 2024 election. The model achieved 98.3% accuracy on the training data, but decreased to 80.8% on the real data, indicating challenges in adapting the model to unstructured and diverse data.

RESEARCH METHODS

This research uses a quantitative approach to conduct sentiment analysis of tweets related to the presidential candidacy. The following are the steps taken in this research :

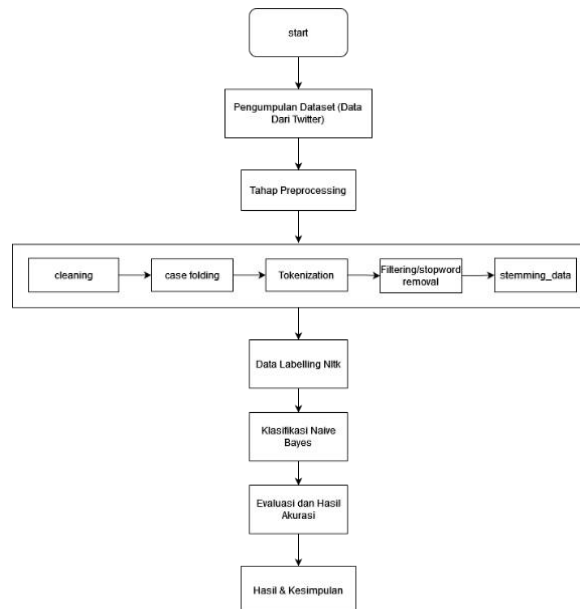


Figure 1. Step Research Flowchart

Dataset Collection

The data for this study was obtained from the social media platform Twitter, with a special focus on tweets related to the topic of Indonesia's presidential candidacy. At this stage, the data taken are Indonesian tweets that discuss the 2024 presidential election. The data collection process is carried out using the Twitter API (Application Programming Interface) which allows systematic retrieval of tweets based on keywords, hashtags, and other parameters relevant to the research.

The tweet selection criteria include :

1. Keywords related to the names of presidential candidates
2. Hashtags popular around the presidential election
3. Time period covering significant moments in the nomination process
4. Tweets in Indonesian language

The total dataset collected was 670 tweets, covering various aspects such as tweet text, creation time, number of retweets, number of likes, and user information. To ensure the diversity and representativeness of the data, sampling was done randomly while taking into account the time distribution and source of the tweets. The data collection process also considered research ethics and user privacy. Sensitive personal information

was removed or anonymised to protect the identity of Twitter users. In addition, only public tweets were included in the research dataset.

Text Preprocessing

After the data retrieval process is complete, then the dataset is carried out in the text preprocessing stage, which is the first step in cleaning the data according to its needs so that text mining can process it further. Text mining algorithms often operate in the same way as data mining techniques (Setiadi 2023). After the data retrieval process is complete, the dataset enters the text preprocessing stage. This stage is a crucial step in cleaning and preparing the data so that it can be processed further by text mining algorithms. Preprocessing aims to improve data quality and consistency, as well as reduce noise that can affect analysis accuracy. In this research, the text preprocessing stage includes the following steps :

1. Cleaning:
 - Removing URLs, mentions (@username), and hashtags (#) from tweets.
 - Removing irrelevant special characters, punctuation marks, and numbers.
 - Removing emoji and emoticons
2. Case Folding:
 - Converts the entire text to lowercase for standardisation.
3. Tokenization:
 - Breaks the text into individual units (tokens), generally words..
4. Filtering/Stopword Removal:
 - Removing common words (stopwords) that have no significant meaning in the analysis, such as 'yang', 'dan', 'di', etc
 - Using a list of Indonesian-specific stopwords.
5. Stemming:
 - Converting words into their base form using the Sastrawi library for Indonesian.
 - Example: 'eat' becomes 'eat', "run" becomes 'run'.

Each of these preprocessing steps has an important role in preparing the data for more accurate sentiment analysis. This process not only cleans up the data, but also reduces the complexity and dimensionality of the data, which in turn can improve the efficiency and effectiveness of the classification algorithms that will be used subsequently.- Example: 'eat' becomes 'eat', "run" becomes 'run'.

Labelling Data The Natural Language Toolkit (NLTK)

NLTK is a suite of natural symbolic and statistical language processing (NLP) libraries and programmes for the English language written in the Python programming language. NLTK is intended to support research and teaching in NLP which includes empirical linguistics, cognitive science, artificial intelligence, information retrieval, and machine learning (Rifano et al. 2020). The majority of tweets had a neutral sentiment, with a total of 666 tweets. Tweets with positive sentiment totalled 599. Tweets with negative sentiments totalled the least, at 68. Overall, negative sentiment.

Naïve Bayes

This strategy uses Bayes' theorem, which was established by Thomas Bayes in the 18th century. Naïve Bayes Classification (NB) is a statistical classification that can be used to predict the probability of class membership. According to Wu and Kumar, the Naïve Bayes algorithm is a well-known classification technique and one of the top 10 algorithms for data mining (Setiadi 2023). Naïve Bayes uses probability theory, a branch of mathematics, to determine the likelihood of possible classifications by analysing the frequency of each classification in the training data.

RESULTS AND DISCUSSION

The following are the results of the research and discussion of the stages in conducting sentiment analysis from data collection, data labelling, text preprocessing, and algorithm implementation.

Dataset Collection

The data used in this study was obtained through twitter with a scraping process. Before the data collection process is carried out, first install the software needed for scraping twitter data, namely python prompt and twitterscraper. The data is stored in .csv form (Nooraeni et al., 2020). which is then used in the preprocessing stage. The data collected in this process was 670 tweets

Full_text

@Andesit14 @henrysubiakto This is long before the presidential nomination bang blm jokowi crack sm bu mega yes maybe the crack is because of this ... Jokowi was disappointed because his son was not nominated by the pdip

@dancersejati07 @officialMKRI Gibran's nomination said the Constitutional Court judges did not find Nepotism Gibran is whose son if

Gibran is not the president's son could he be nominated for the vice presidency so clearly this is nepotism the Constitutional Court judges violated their oath of office they are democracy killers. #GerakanAntiJokowi #GerakanAntiJokowi
MPR Speaker Bambang Soesatyo encouraged the revision of the Election Law to make a number of improvements and refinements. improvements include the parliamentary threshold election system to the threshold for nominating the president and vice president. — #HabisPemiluMariBersatu #ChampionsLeague https://t.co/T6dI8ZkQhz

Table 1. Twitter API Data

number of improvements and refinements. improvements include the parliamentary threshold election system to the threshold for nominating a president and vice president. — #HabisPemiluMariBersatu #ChampionsLeague https://t.co/T6dI8ZkQhz	number of improvements and refinements. improvements include the electoral system parliamentary threshold to the threshold for nominating the president and vice president. — #habispemilumaribersatu #championsleague
---	--

Table 2. Case Folding

Data Preprocessing

The next process is preprocessing. This stage has the benefit of producing good input data for the training and testing process. The preprocessing stage is a stage for selecting words in tweets so as to produce words that contain sentiment by removing words that are not needed (Faisal Rahutomo et al., 2018)

Full_text	Case_folding
Abah's congratulations to Pak Prabowo for his victory in the 2024 presidential election after the Constitutional Court ruled with 3 judges granting 5 judges rejecting the presidential election lawsuit by 01&03 yesterday. Congratulations to Pak Prabowo for fulfilling his dream of becoming president after 2 candidacies! https://t.co/dutu3y3wVa	congratulations to pak prabowo for his victory in the 2024 presidential election after the mk ruled with 3 judges granting 5 judges rejecting in the presidential election lawsuit by 01&03 yesterday. congratulations to pak prabowo for fulfilling his dream to become president after 2 candidacies!
President Jokowi's blessing or Gibran's candidacy is thought to have a major influence on the Prabowo-Gibran victory. https://t.co/4QFMyZPGFY #Opinion #AdadiKompas	president jokowi's blessing or gibran's candidacy is thought to have had a major influence on the prabowo-gibran victory. #opini #adadikompas
MPR Speaker Bambang Soesatyo encouraged the revision of the Election Law to make a	mpr chairman bambang soesatyo encouraged the revision of the election law to make a

This table shows the comparison between the original text (full text) and the results of the case folding process in the preprocessing stage of sentiment analysis. The left column shows the original text of some social media posts related to Indonesian political issues, while the right column shows the results after case folding is applied. This process converts all letters to lowercase, but retains punctuation marks, numbers, and special symbols such as hashtags. Case folding is an important step in text standardisation to prepare the data for further sentiment analysis.

Full_text	Tokenization
Abah's congratulations to Pak Prabowo for his victory in the 2024 presidential election after the Constitutional Court ruled with 3 judges granting 5 judges rejecting the presidential election lawsuit by 01&03 yesterday. Congratulations to Pak Prabowo for fulfilling his dream of becoming president after 2 candidacies! https://t.co/dutu3y3wVa	congratulations,abah,t o,pak,prabowo,on,his, victory,in,the,presiden tial,election,2024,after ,the,mk,ruled,with,3,ju dges,granting,5,judges ,rejecting,in,the,presid ential,election,lawsuit, by,01&03,yesterd ay., congratulations, to, pak, prabowo, fulfilled, also, the goal, to be, president, after, 2 times, nomination!
President Jokowi's blessing or Gibran's candidacy is thought to have a major influence on the Prabowo-Gibran victory. https://t.co/4QFMyZPGFY #Opinion #AdadiKompas	blessing, president, jokowi, or, nomination, gibran, allegedly, a big influence, on, victory, prabowo-gibran.,

Nasdem may be the first non-coalition party to confirm that it will support and cooperate with the President-elect. I'm not surprised because the big scenario has been read before the registration and even during the nomination.	#opinion,#adadiko mpas nasdem,may,be,the, non,coalition,party,t hat,is,the,first,to,co nfirm,that,it,will,su pport,and,cooperate ,with,the,president,e lect.,no,surprise,bec ause,the,big,scenari o,has,already,been,r ead,before,registrati on,even,during,the, nomination.
--	--

Table 3. Step Tokenization

This table shows the tokenisation process in preprocessing text data for sentiment analysis. The table has two columns: 'Full text' (original text) and 'Tokenisation' (tokenisation result). Tokenisation breaks the text into small units or 'tokens', generally individual words, by separating words and punctuation using commas. This process is important to prepare the text data for further analyses such as word frequency calculation or the application of machine learning algorithms. Example shown using text related to Indonesian political issues from social media.

Full_text	Tokenization
@dancersejati07 @officialMKRI Gibran's nomination said the Constitutional Court judges did not find Nepotism Gibran is whose son if Gibran is not the president's son could he be nominated for vice president so clearly this is nepotism the Constitutional Court judges violated their oath of office they are democracy killers. #GerakanAntiJokowi #GerakanAntiJokowi	@dancersejati07,@offi cialmkri,pencalonan,gi bran,hakim,mk,tdk,me nemukan,nepotisme,gi bran,anak,klo,gibran,b kn,anak,presiden,calo nkan,cawapres,nepoti sme,hakim,mk,melang gar,sumpah,jabatan,m rk,pelaku,pembunuh,d emokrasi.,#gerakanan tijokowi,#gerakanantij okowi
@Andesit14 @henrysubiakto This is long before the presidential nomination bang before jokowi cracked sm bu mega yes	@andesit14,@henrys ubiakto,sblm,pencalona n,presiden,bang,sblm,j okowi,retak,sm,bu,me ga,ya,retaknya,gara2,i ni.,jokowi,kecewa,krn,

maybe the crack is because of this ... Jokowi is disappointed because his son is not nominated by the pdip	anaknya,ngga,dicalonk an,olh,pdip
@inijoejoe @prabowo @gibran_tweet @beacukaiRI @prastow All supporting parties can be angry if the open recruitment selection is tight. Because the supporting parties ask for a ration of seats if there is no ration the party will be cranky why the party supports the nomination of the president and vice president. This is where bad ministerial performance starts	@inijoejoe,@prabowo, @gibran_tweet,@beac ukairi,@prastow,parta i,pendukung,marah,op en,rekrutmen,seleksi, ketat,partai,pendukun g,jatah,kursi,gak,jatah, partai,ngambek,partai, mendukung,pencalona n,presiden,wakil,presi den.,buruk,kinerja,me nteri

Table 4. Filtering / Stopword Removal

Full_text	Steaming_data
@ommi_siregar That's why the presidential nomination system must be changed, freeing all Indonesians to register and then strictly selected by the DPR with agreed rules and laws, there will be no more imaging and the elected president will be easier to monitor.	@ommi_siregar, system, nomination, president, changed, free, people, indonesia, register, selection, strict, parliament, rules, laws, agreed, no more, imaging, similar, president, elected, easy, supervised
Nasdem may be the first non-coalition party to confirm that it will support and work with the President-elect. I'm not surprised because the big scenario has been read before registration and even during the nomination.	snasdem,a,non,coaliti on,party,confirmed,its ,support,for,co- operation,with,the,pr esident, elect.,no,surprise,the, scenario,has,already, been,read,before,the,r egistration,of,nominat ions
Chairman of Commission II of the House of Representatives Ahmad Doli Kurnia said that the presidential nomination threshold will be one of the targets of the	chairman, commission,ii,dpr,ri,a hmad,doli,kurnia,calls ,threshold,nominatio n,for,president,one,ta rget,of,revision,of,elec tion,law.

revision of the Election Law. https://t.co/wdYuLMo9zc	
--	--

Table 5. Steaming Data

Data Labelling Using NLTK Library

Index	stemming_data	sentiment_score	sentiment
0	andesi14 herryudaito olim calon presiden bang olim jkowi retak sm bu mega ya retak gara2 ni jkowi kecewa lim anak ngga calon oih pdp	0.0	negatif
1	dancesegat07 officialmisi calon ghan hakim mik tdk temu negosiasi ghan anak klo ghan bin anak presiden calon caupres negosiasi hakim mik langgar sumpah jabat mik laku bunuh demokrasi gesakanantipjokowi gerakanantipjokowi	0.0	negatif
2	nipepe prabowo ghan tweet beacikain praslow partai dukung marah open rekutimen seleksi ketat partai dukung jatah kurir gak jatah partai ngambek partai dukung calon presiden wakil presiden buruk kerja menten	0.0	negatif
3	ketua mpr bambang sonesatyo dorong laku revisi uu mlu utk baik sempurna baik sistem mlu ambang batas parlemen ambang batas calon presiden wapres habasemulmabesatu champonsileague	0.0	negatif
4	putus mik hasil evaluasi mlu tund sempurna uu mlu datang kait sistem mlu ambang batas parlemen ambang batas calon presiden-wakil presiden	0.0	negatif
5	legolawea12 agnes nur4007 realfidmuli halo prabowo cull tantes parson 01 sponsor cull calon gubernur diti tum pdp pengusung capres 03 gandong beliau pili presiden wakil lalu	0.0	negatif
6	dednyupalaka dunia runtuh dipungiri ghan folo calon kondisi mik kpu proses benar beliau anak presiden anak ketua kpu anak menten di proses masalah fair	0.3102	positif
7	vanc10aczz jkowi abuk anak anak jd presiden lupa ralyat sendi	0.0	negatif
8	doankinto layak jd presiden pdp dahulu calon presiden suara pdp yg besar balik calon pdp walikota gubernur presiden si main solo	0.0	negatif
9	murtadhaone1 mantab merantu presiden jkowi officialmisi negosiasi dlm calon ghan tweet langit bobhyneation luar biasa	0.0	negatif
10	dednyupalaka kiamat hapus calon ghan tweet lepas campur tangan sang paman mik yg bukti langgar etika anak satu nya presiden yg jabat tarik bulu shydyhonyo tu	0.0	negatif
11	low umkm4 jkowi pemah presiden buka terang benderang umum calon dgn sebut nama	0.6389	positif
12	bedah bulu rekonstruksi ambang batas calon presiden ada hukum universitas sulawesi barat fyo hukum ltn sorot law	0.0	negatif
13	mio071i iam zell calon prabowo subianto presiden nggakt tepat pdp anggap gara2 bangkit jkowi jkowi megawati nggakt diti	0.0	negatif
14	baca analisisbudaya anel heryanto judul pria idem tulis restu presiden jkowi calon ghan duga pengaruh menang prabowo-ghan adalokompas	0.0	negatif

Figure 2. Sentiment Stages

Sentiment Analysis Results

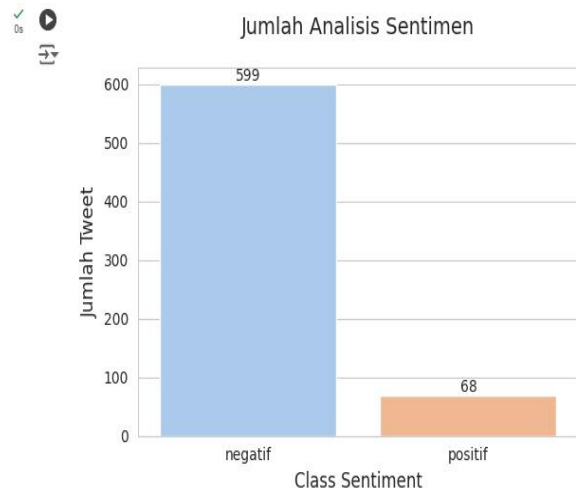


Figure 3. Sentiment Visualisation

CONCLUSIONS AND SUGGESTIONS

Conclusion

This research shows that the Naïve Bayes algorithm is effective in analysing public sentiment towards the 2024 presidential candidacy on social media. With high accuracy and supported by SMOTE technique, the model is able to overcome data imbalance and classify opinions well. This finding underscores the importance of machine learning-based sentiment analysis in understanding the dynamics of public opinion digitally.

Suggestion

Future research is recommended to use larger and more diverse datasets, as well as compare the Naïve Bayes algorithm with other methods such as SVM or LSTM. In addition, combining sentiment analysis with interactive data visualisation can help understand public opinion trends in real-time.

REFERENCES

- Al-Areef, M. H., & Saputra, K. (2023). Analisis Sentimen Pengguna Twitter Mengenai Calon Presiden Indonesia Tahun 2024 Menggunakan Algoritma LSTM. *Agustus*, 22, 270–279. <https://ojs.trigunadharma.ac.id/index.php/jis/index>
- Aurelly Claudia Budianto, T., Fatoni, H., Ayu Syaharani, M., Rozikin Informatika, C., Singaperbangsa Karawang Jl HSRonggo Waluyo, U., & Timur, T. (2024). ANALISIS SENTIMEN PENGUMUMAN HASIL PEMILU 2024 DI SOSIAL MEDIA X MENGGUNAKAN KNN DAN NAÏVE BAYES CLASSIFIER. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 5).
- Duei Putri, D., Nama, G. F., & Sulistiono, W. E. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika Dan Teknik Elektro Terapan*, 10(1). <https://doi.org/10.23960/jitet.v10i1.2262>
- Faisal Rahutomo, Saputra, P. Y., & Fidyawan, M. A. (2018). Implementasi Twitter Sentiment Analysis Untuk Review Film Menggunakan Algoritma Support Vector Machine. *Jurnal Informatika Polinema*, 4(2), 93. <https://doi.org/10.33795/jip.v4i2.152>
- Fauza Majbur, R., Yanuar, F., & Devianto, D. (2024). PENERAPAN METODE NAÏVE BAYES CLASSIFIER DALAM MENGANALISIS SENTIMEN PADA MEDIA SOSIAL X TERHADAP PILPRES 2024 DI INDONESIA. 5(3). <https://doi.org/10.46306/lb.v4i1.794>
- Fikri Haikal, M., & Indra, J. (2024). Analisis Sentimen Bakal Calon Presiden Indonesia 2024 Dengan Algoritma Naïve Bayes.
- Gunawan, Y., & Purnama, I. (n.d.). Sentiment Analysis Of Twitter Towards The 2024 Indonesian Presidential Candidates Using The Naïve Bayes Algorithms. In *International Journal of Science*. <http://ijstm.inarah.co.id>953

- Haque, S., Eberhart, Z., Bansal, A., & McMillan, C. (2022). Semantic Similarity Metrics for Evaluating Source Code Summarization. *IEEE International Conference on Program Comprehension*, 2022-March, 36–47. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>
- Luthfiansyah Dan, R., Wasito, B., Program, S., & Sistem, I. (2023). *ANALISIS SENTIMEN TERHADAP PARA KANDIDAT PRESIDEN 2024 BERDASARKAN NETIZEN PENGGUNA TWITTER DENGAN METODE DATA MINING DAN TEXT MINING*.
- Nugroho, D. S., Hanif, I. F., Hasbi, M. A., Fredianto, F., Saputra, A. M., & Zildjian, R. (2024). Analisis Sentimen Dugaan Pelanggaran Pemilu 2024 Berdasarkan Tweet Menggunakan Algoritma Naïve Bayes Classifier. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 1169–1176. <https://doi.org/10.57152/malcom.v4i3.1496>
- Permana, H., Herry Chrisnanto, Y., Ashaury Informatika, H., Jenderal Achmad Yani Cimahi Jl Terusan Jend Sudirman, U., Cimahi Sel, K., Cimahi, K., & Barat, J. (2023). ANALISIS SENTIMEN TERHADAP BAKAL CALON PRESIDEN 2024 DENGAN ALGORITMA MULTINOMIAL NAÏVE BAYES DAN OVERSAMPLING SMOTE. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 5).
- Prayoga Siswono, A., Fauzi, S., Zalva Surya Hermawan, A., Riyandi, A., Panjaitan No, J. DI, Kidul, P., Purwokerto Selatan, K., & Banyumas, K. (2024a). *Analisis Sentimen Pelantikan Presiden Indonesia 2024 Menggunakan Model Klasifikasi dan Algoritma Naive Bayes* (Vol. 4, Issue 1).
- Prayoga Siswono, A., Fauzi, S., Zalva Surya Hermawan, A., Riyandi, A., Panjaitan No, J. DI, Kidul, P., Purwokerto Selatan, K., & Banyumas, K. (2024b). *Analisis Sentimen Pelantikan Presiden Indonesia 2024 Menggunakan Model Klasifikasi dan Algoritma Naive Bayes* (Vol. 4, Issue 1).
- Priansyah, E., & Sutabri, T. (2024). IJM: Indonesian Journal of Multidisciplinary Analisis Sentimen Berbasis Naïve Bayes Pada Media Sosial Twitter Terhadap Hasil Pemilu Indonesia 2024. *IJM: Indonesian Journal of Multidisciplinary*, 2. <https://journal.csspublishing/index.php/ijm>
- Putri Nardilasari, A., Lia Hananto, A., Shofia Hilabi, S., & Priyatna, B. (2023). *Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM Pada Media Sosial Twitter* (Vol. 7, Issue 1).
- Subowo, M. H., & Alzami, F. (2024). *Using 2024 election twitter data, sentiment analysis based on TF-IDF and Naïve Bayes*.
- Vindua, R., & Zailani, A. U. (2023). *Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python*.
- Yin, H., Yang, S., & Li, J. (2020). *Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media*. <http://arxiv.org/abs/2007.02304>
- Yudha Patria, W., Harry Gunawan, P., & Aquarini, N. (2025). Public Political Sentiment Post 2024 Presidential Election: Comparison of Naïve Bayes and Support Vector Machine. *Technology and Science (BITS)*, 6(4). <https://doi.org/10.47065/bits.v6i4.6734>
- Yusuf Ramadhan Nasution, Suhardi Suhardi, & Ilham Hafiz Satrio. (2024). Penerapan Algoritma Klasifikasi Naïve Bayes Untuk Analisis Sentimen Tentang Pemilu 2024. *Elkom: Jurnal Elektronika Dan Komputer*, 17(2), 495–502. <https://doi.org/10.51903/elkom.v17i2.2053>