JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

SENTIMENT ANALYSIS OF PUBLIC OPINION ON SOCIAL MEDIA X TOWARDS ETHNIC ROHINGYA IN INDONESIA

Ayu Amalia[1], Damayanti[2*]
[1]Information Systems, Universitas Teknokrat Indonesia
email: *damayanti@teknokrat.ac.id

Abstract: Rohingya refugees continue to arrive in Aceh by sea by boat. Based on data from the United Nations High Commissioner for Refugees (UNCHR), as of December 10 2023, 1,543 Rohingya refugees had landed in Aceh since mid-November 2023. The increasing number of refugees arriving has caused resistance from local residents. This rejection was the result of the bad experiences of Acehnese people with Rohingya refugees. The main problem of this research is that analyzing public opinion on Rohingya ethnicity in Indonesia is still done manually by looking at tweets one by one. The solution to overcome this is to analyze opinions using data crawling with the Naïve Bayes algorithm. The purpose of this research is to determine public opinion on Twitter media regarding refugee refugees. The method used in this research is the Naïve Bayes algorithm method. The results of the research show that the Naïve Bayes algorithm can classify public opinion sentiment on Twitter social media towards the Rohingya Ethnic in Indonesia into positive sentiment. and negative with a total accuracy of 70%. So, the words "Rohingya Ethnicity in Indonesia" tend to be accepted by the X community with the arrival of Rohingya refugees in Indonesia.

Keywords: rohingya; twitter; naïve bayes; opinion

Abstrak: Pengungsi Rohingya terus berdatangan ke Aceh melalui jalur laut dengan menggunakan perahu. Berdasarkan hasil data United Nations High Commissioner for Refugees (UNCHR), per 10 Desember 2023 sebanyak 1.543 pengungsi Rohingya datang ke dalam wilayah Aceh. Meningkatnya jumlah pengungsi yang datang menimbulkan perlawanan dari warga setempat. Penolakan ini imbas dari pengalaman buruk warga Aceh terhadap pengungsi Rohingya. Permasalahan utama penelitian ini yaitu menganalisis opini publik terhadap Etnis Rohingya di Indonesia masih secara manual dengan melihat tweet satu persatu. Solusi mengatasi hal tersebut maka analisis opini menggunakan crawling data dengan algoritma Naïve Bayes. Tujuan penelitian ini untuk mengetahui opini publik pada media twitter terkait pengungsi Rohingya. Metode yang digunakan dalam penelitian ini yaitu metode algoritma Naïve Bayes. Hasil penelitian menunjukan bahwa pada algoritma Naïve Bayes dapat mengklasifikasikan sentimen dengan total akurasi 70%. Maka “Etnis Rohingya di Indonesia” cenderung dapat diterima oleh masyarakat X dengan datangnya pengungsi Rohingya di Indonesia.

Kata kunci: rohingnya; twitter; naïve bayes; opini

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

INTRODUCTION

In today's modern era, the internet has become a common thing for people around the world, especially social media. Social media is widely used as a means to communicate. Social media is a place for people to convey their wishes, ideas, and even criticism. Social media can be accessed via cell phones or computers with internet networks [1].

Indonesia is included in the group of countries with people who use the seventh largest X social media in the world ranking. This is based on the quantity of X application users in Indonesia, which was 13.8 million per year in 2020 and is expected to reach 17.8 million in 2025 [2]. Social media is a place where people can express their wishes about a problem through tweets. This tweet can be a source of data that can be used to collect public opinion on a problem.

Public opinion is one of the data sources in opinion mining. Opinion mining, commonly translated as sentiment analysis, is one of the text mining implementations that is easy to understand, extract, and process sentiment in opinion sentences [3].

Sentiment analysisis is natural languange processing [5]. Sentiment analysis is often done using the Naïve Bayes algorithm.

Sentiment analysis is used to analyze the opinions of X social media users on topics or issues that are currently being discussed. The Rohingya ethnic issue in Indonesia is currently one of the topics that is widely discussed by X social media users in Indonesia. This stems from the humanitarian crisis in Rakhine, Myanmar, becoming an alarming and complex problem. In response, the Indonesian people have held several demonstrations to express their concern about the conflict in Myanmar. People put pressure on the government to participate in resolving conflicts that affect the oppressed Rohingya group. In 2015, hundreds of Rohingya refugees arrived through the waters of Aceh. During that period, Rohingya refugees continued to arrive in Aceh by sea, using boats. Based on data from the United Nations High Commissioner for Refugees (UNCHR), as of December 10, 2023, as many as 1,543 Rohingya refugees had arrived in the Aceh region since mid-November 2023. The increasing number of refugees who came caused resistance from local residents. This was exacerbated by the spread of hoaxes and hate speech against Rohingya refugees through social media. This not only worsened the view of the Acehnese people but also the view of the Indonesian people towards the Rohingya [7].

This phenomenon also encompasses public sentiment towards the Rohingya ethnic group in Indonesia. Several studies have provided valuable insights into the complexity of factors influencing public opinion towards the Rohingya ethnic group and its impact on social and political dynamics in Asean [13] . For instance, research by Putra in 2021 highlighted how social media influences public opinion towards the Rohingya ethnic group, [14] while a study by Silviya clarified the relationship between political dynamics and public perception towards the Rohingya ethnic group [15].

Furthermore, research by Tusriyanto in 2023 investigated how interactions among ethnic groups in Indonesia affect public opinion towards Rohingya refugees [16], while research by Kurniati in 2018 explored the gender

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

impact on public opinion towards the Rohingya ethnic group [17]. Findings from these studies indicate that public perceptions towards the Rohingya ethnic group are influenced by various factors, including media, politics, religion, culture, inter-ethnic interactions, and gender factors [18].

The main problem of this research is that analyzing public opinion on Rohingya ethnicity in Indonesia is still done manually by looking at tweets one by one. The solution to overcome this is to analyze opinions using data crawling with the Naïve Bayes algorithm. This research was conducted to find out public opinion on Rohingya ethnicity in Indonesia so that the data obtained can be analyzed by grouping them in a class.

METHOD

This stage explains how this research was conducted. Text mining is the collection of structured, unstructured, or semi-structured data, such as emails, text documents, and HTML files. So text mining is considered a much better solution [8]. Sentiment analysis is a group of text polarities present in documents, sentences, or opinions [9]. Therefore, this study aims to analyze the review of public opinion comments in the form of tweets about the Rohingya ethnicity in Indonesia on X social media.

This research uses the multinomial Naïve Bayes algorithm classification method. In general, multinomial Naïve Bayes has better performance in classifying documents, especially on large amounts of data [10]. The completion of this research involves several stages. The research stages can be seen in Image 1.

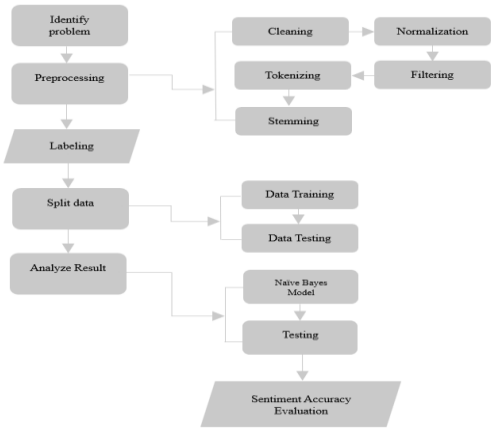

Image 1. Research Stages

The research stages can be explained as follows:

Problem Identification

Observe and find problems that occur in public opinion tweets about the Rohingya ethnicity in Indonesia on X social media from December 10, 2023, to January 15, 2024.

Data Collection

This study uses data from as many as 3555 public opinion tweets obtained through the X application and then processed so that it can become material that is ready to be tested.

Data Preprocessing

Preprocessing is one of the important stages for data in the text mining process. The purpose of preprocessing is to convert data into a format that is easy to understand and also organize the data format for the better.

There are several stages that are passed in the preprocessing process, including: (1) Cleaning is used to remove some characters that are not included in the alphabet. For example, removing emoticons, removing numeric characters, hastags (#), mentions (@), and removing URLs and links in comments. (2)

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Normalization is converting the original data into a format that allows data processing to be efficient., (3) Filtering is the process of selecting words from a token. The algorithms used are stopwords, or removing words that are not needed., (4) Tokenizing is separating documents into word chunks., (5) Stemming is changing the base word of a word that has affixes., (6) Split Data.

Data will be separated into two parts, namely, training data and testing data. Analyze Results. In this stage, the modeling process is carried out using the multinomial Naïve Bayes algorithm, which is then implemented into the testing data. So that it can produce prediction evaluation values and the accuracy of the sentiment analysis modeling process.

RESULT AND DISCUSSION

Data Crawling

The data crawling process is the process of collecting data in the form of tweets processed through Google Collaboratory, which are then stored in .csv. format.

	full_text	username
0	@sgagovind @neohistoria_id @geelvinkbaai Artin...	yumiriinchan
1	si rohingya ngumpetin sepatu gwa!!	pinoctvrnal
2	Selasa (09/01) Kantor Imigrasi Dumai melakukan...	imigrasi_dumai
3	@laurquers @Prnchybrid Iya cuma pgn komen di p...	pistoletchaud
4	@angewwie Mau dihibahkan buat nampung pengungs...	babijoged
...	...	...
3550	@neohistoria_id @m_marshel Oh ini bazzer unchr ya	Kav_ya13
3551	@kegblgnunfaedh UNCHR cek bio aku yahh	sipalingchindo
3552	@UNHCRIndo @UNinIndonesia Pasal pasal mulu lo ...	tingkerbong
3553	@melanofita93705 @Greschinov Di samping isrewe...	whoyouarevi
3554	@sosmedkeras Anehnya ada bbrpa akun gede indo ...	Aaaa87223662624

3555 rows × 2 columns

Image 2. Crawling Data

From the data has been obtained through the crawling process only full_text will be processed.

Data Cleaning

In the data cleaning process, data cleaning is carried out by removing emojis, links, special characters, and numbers. Then do the case folding process, which involves converting all data into lowercase letters.

full_text	cleaning	case_folding
@sgagovind @neohistoria_id @geelvinkbaai Artin...	Artinya kewarganegaraan mereka bisa dimanipula...	artinya kewarganegaraan mereka bisa dimanipula...
si rohingya ngumpetin sepatu gwa!!	si rohingya ngumpetin sepatu gwa	si rohingya ngumpetin sepatu gwa
Selasa (09/01) Kantor Imigrasi Dumai melakukan...	Selasa Kantor Imigrasi Dumai melakukan peminda...	selasa kantor imigrasi dumai melakukan peminda...
@laurquers @Prnchybrid Iya cuma pgn komen di p...	Iya cuma pgn komen di persamaan orang Rohingny...	iya cuma pgn komen di persamaan orang rohingny...
@angewwie Mau dihibahkan buat nampung pengungs...	Mau dihibahkan buat nampung pengungsi Rohingya...	mau dihibahkan buat nampung pengungsi rohingya...

Image 3. Data Cleaning

Text Preprocessing

There are several stages in text preprocessing, including the following: Normalization is converting the original data into a format that allows efficient data processing. Tokenizing is the process of cutting text into smaller parts, called tokens. Stopword removal is the process of removing words that are not important in text analysis. For example, "yang", "saya", "iya", "buat", and others. Stemming is the process of converting word forms into basic words. For example, "dihibahkan" becomes "hibah".

The following are the results of the four processes in text preprocessing.

normalization	tokenizing	stopwords	stemming
artinya kewarganegaraan mereka bisa dimanipula...	[artinya, kewarganegaraan, mereka, bisa, diman...	[artinya, kewarganegaraan, dimanipulasi, bangl...	arti kewarganegaraan manipulasi bangladesh aku...
si rohingya ngumpetin sepatu saya	[si, rohingya, ngumpetin, sepatu, saya]	[si, rohingya, ngumpetin, sepatu]	si rohingya ngumpetin sepatu
selasa kantor imigrasi dumai melakukan peminda...	[selasa, kantor, imigrasi, dumai, melakukan, p...	[selasa, kantor, imigrasi, dumai, melakukan, p...	selasa kantor imigrasi dumai laku pindah wna d...
iya hanya ingin komen di persamaan orang rohin...	[iya, hanya, ingin, komen, di, persamaan, oran...	[komen, persamaan, orang, rohingnya, israjal, ...	komen sama orang rohingnya israjal kasus rohin...
mau dihibahkan buat nampung pengungsi rohingya...	[mau, dihibahkan, buat, nampung, pengungsi, ro...	[mau, dihibahkan, buat, nampung, pengungsi, ro...	mau hibah buat nampung ungsi rohingya kak

Image 4. Text Preprocessing

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Data Labeling

The data labeling process is carried out based on the full_text attribute, which produces positive and negative labels based on tweet data.

full_text	Sentiments
anies mention konflik papua isu masuk ungsi ro...	Negatif
besar mungkin ungsi rohingya susup sama oknum ...	Negatif
sidik satreskrim polresta banda aceh periksa j...	Negatif
internet mudah cari informasi kamu liat 1 site...	Positif
ttp ga etis sih bilang rohingya ky israajal be...	Positif

Image 5. Data Labeling

Label Visualization

At the label visualization stage, it is carried out on data that has been processed previously. The comparison results of positive sentiment analysis and negative sentiment analysis can be seen in the following image.

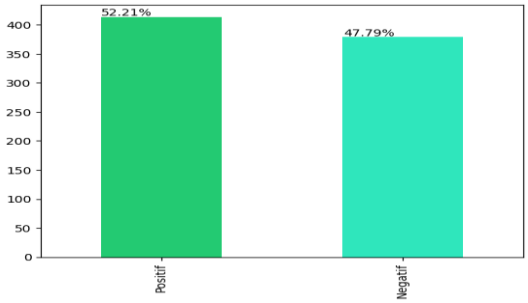

Image 6. Label Visualization

Wordcloud Sentiment Results

Wordcloud is a visualization that displays words from text with a larger font size for words that appear frequently. One of the functions of Wordcloud is to evaluate the content of the text and find the main topics discussed. The visualization of positive and negative sentiment Wordcloud can be seen in the following image.

Image 7. Wordcloud Positive

Image 8. Wordcloud Negative

Build Model

The multinomial Naive Bayes algorithm is one of the probabilistic learning methods based on the Bayes theorem used in Natural Language Processing (NLP). This algorithm works on the concept of term frequency, which means how many times a word appears in a document [11]. MultinomialNB works by calculating the probability of the occurrence of words in the text that have been labeled positive and negative. The following are the accuracy results of the model that has been built with the MultimomialNB algorithm.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Akurasi Model Naive Bayes: 0.7058823529411765

Laporan Klasifikasi :

	precision	recall	f1-score	support
Negatif	0.62	0.95	0.75	56
Positif	0.91	0.49	0.64	63
accuracy			0.71	119
macro avg	0.77	0.72	0.70	119
weighted avg	0.78	0.71	0.69	119

Image 9. Accuracy Result

Evaluation

Confusion Matrix is a visual evaluation tool used in machine learning. The columns in Confusion Matrix represent the predicted class results and the actual rows. A series of algorithm performance measures can be defined using Confusion Matrix [12]. This evaluation can be used as a reference of the classification performance used.

Table 1. Confusion Matrix

Confusion Matrix		Prediksi	
		Positif	Negatif
Aktual	Positif	TP	FN
	Negatif	FP	TN

Description:
True Positive (TP): A true positive outcome is classified as a positive class.
True Negative (TN): A true negative outcome is classified as a negative class.
False Positive (FP): A negative outcome that is incorrectly classified as a positive class.
False Negative (FN): A positive outcome that is misclassified as a negative class.

CONCLUSION

Based on the results of the research conducted, it is known that the Naïve Bayes algorithm can classify the sentiment of public opinion on X social media towards the Rohingya ethnicity in Indonesia into positive and negative sentiments with a total accuracy of 70%. So the words "Ethnic Rohingya in Indonesia" tend to be acceptable to the public with the arrival of Rohingya refugees in Indonesia.

BIBLIOGRAPHY

[1] D. General and D. Available, "Media pembelajaran IPS pasca pandemi Covid-19 Yayan Sudrajat Universitas Indraprasta PGRI," vol. 8, no. July, pp. 29–34, 2022.

[2] F. A. Girnanfa and A. Susilo, "Studi Dramaturgi Pengelolaan Kesan Melalui Twitter Sebagai Sarana Eksistensi Diri Mahasiswa di Jakarta," *J. New Media Commun.*, vol. 1, no. 1, pp. 58–73, 2022, doi: 10.55985/jnmc.v1i1.2.

[3] B. Edgar Maulana Thoriq,Rahayudi and D. E. Ratnawati, "Analisis Sentimen Opini Publik pada Media Sosial Twitter terhadap Vaksin Covid-19 menggunakan Algoritma Support Vector Machine dan Term Frequency-Inverse Document Frequency," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 12, pp. 5349–5355, 2021.

[4] A. D. Adyatma, L. Afuan, and E. Maryanto, "The Effect of Unigram and Biagram in the Naïve Bayes Multinomial for Analyzing of Comment Sentiment of Gojek Application in Google Play Store," vol. 4, no. 6, pp. 1535–1540, 2025.

[5] Ardiansyah, Adika Sri Widagdo,

Krisna Nuresa Qodri, F. E. N. Saputro, and Nisrina Akbar Rizky P, "Analisis sentimen terhadap pelayanan Kesehatan berdasarkan ulasan Google Maps menggunakan BERT," *J. Fasilkom*, vol. 13, no. 02, pp. 326–333, 2023, doi: 10.37859/jf.v13i02.5170.

[6] M. N. Akbar and Nirwana Samrin, "Analisis Sentimen Komentar Pengguna Aplikasi Threads Pada Google Playstore Menggunakan Algoritma Multinominal Naive Bayes Classfier," *AGENTS J. Artif. Intell. Data Sci.*, vol. 3, no. 2, pp. 21–29, 2023, doi: 10.24252/jagti.v3i2.67.

[7] C. A. Sopamena, "Pengungsi Rohingya Dan Potensi Konflik & Kemajemukan Horizontal Di Aceh," *Caraka Prabu J. Ilmu Pemerintah.*, vol. 7, no. 2, pp. 85–115, 2023.

[8] A. Hermawan, I. Jowensen, J. Junaedi, and Edy, "Implementasi Text-Mining untuk Analisis Sentimen pada Twitter dengan Algoritma Support Vector Machine," *JST (Jurnal Sains dan Teknol.*, vol. 12, no. 1, pp. 129–137, 2023, doi: 10.23887/jstundiksha.v12i1.52358 .

[9] E. T. Handayani and A. Sulistiyawati, "Analisis Sentimen Respon Masyarakat Terhadap Kabar Harian Covid-19 Pada Twitter Kementerian Kesehatan Dengan Metode Klasifikasi Naive Bayes," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 3, pp. 32–37, 2021, [Online]. Available: http://jim.teknokrat.ac.id/index.php/JTSI

[10] M. Sahid, "Penggunaan Media Sosial Dalam Peningkatan Pendaftar Mahasiswa Baru," *J. Inov. Penelit.*, vol. 3, no. 8, pp. 7417–7428, 2023.

[11] Darussalam and G. Arief, "Jurnal Resti," *Resti*, vol. 1, no. 1, pp. 19–25, 2018.

[12] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf. Sci. (Ny).*, vol. 507, pp. 772–794, 2020, doi: 10.1016/j.ins.2019.06.064.

[13] Flanders,S., & Nungsaru,M. (2022) A textual analysis of the effect of short them voluntering on attitudes toward refugees in Malaysia. *The Social Journal, 2022, 1-23*

[14] Putra,I.E., Selvanathan, H. P., Mashuri, A., & Moontiel, C. K. (2021). Aungg Suuu Kyi's defensive denial of the rohingya massacre: A rhetorical analysis of denial and positiv- image construction. *Journal of Social and Poliical Psychology*, 9 (2), 353-369

[15] silvya Sari, D. (2018). Indonesian government policy on Rohingya refugees. Andalas Journal of International Studies (AJIS), 7(1), 1-13

[16] Tusriyanto, T., Murthada, M., Nuswantor, P., & Maulinda, R (2023). Pengaruh Persepsi Masyarakat, Identitas Budaya dan Tingkat Akulturasi pada Proses Adaptasi Migran Rohingya di Lingkungan Baru Aceh, Indonesia. *Sanskara Ilmu Sosial dan Humaniora,* 1 (01), 10-20

[17] Kurniawati N. Q. (2018). Persepsi Paguyuban Antar Umat Beragama Penghayat Kepercayaan (Paub-Pk) Kediri Terhadap Konflik Rohingya Myanmar (Analisis Deskriptif Aktivis Paguyuban Antar Umat Be-

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. X No 3, Juni 2024, hlm. 443 - 450 ISSN 2550- 020 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v10i3.3079
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

ragama Pengahayat Kepercayaan Kota Kediri) (Doctoral dissertation, IAIN Kediri).

[18] Mamis, S., Rustan, A. S., Arnus, S. H., & Basri, H. (2023). Framing Media Dan Dinamika Opini Publik TerkaitPengungsi Rohingya Di Indonesia: Perspektif UNHCR Dan Otoritas Aceh Dalam Sorotan Kritis. Wardah, 24(2), 137-163