Analyzing the Indonesian sentiment to rohingya refugees using IndoBERT model

M. Zainal Arifin [a,1,*], Sandy Yunan Maulana [a,2], Agustinus Noertjahyana [b,3], Asghaiyer Mohamed Asghaiyer [c,4]

[a] Informatic Departmen, State University of Malang, Malang, Jawa Timur, Indonesia
[b] Informatic Departmen, Petra University, Jawa Timur, Indonesia
[c] Bani Waled School, Libya
[1] arifin.mzainal@um.ac.id; [2] sandy@um.ac.id; [3] agust@petra.ac.id; [4] arifin.mzainal@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received June 17, 2024
Revised June 28, 2024
Accepted July 13, 2024

Keywords
Sentiment analysis
BERT
IndoBERT
Rohingya
Confusion matrix

ABSTRACT

This study aims to analyze public sentiments towards Rohingya refugees in Indonesia using the IndoBERT model. We collected sentiment data from social media platforms and news articles, followed by preprocessing techniques including tokenization, cleaning, case folding, stemming, and filtering. Sentiment labels were assigned using the InSet lexicon, and the IndoBERT model was trained with these labeled data. Our findings reveal that the predominant sentiment is negative, with 65% of the sentiments classified as negative, 20% as neutral, and 15% as positive. The model demonstrated robust performance with an accuracy of 87%, precision of 85%, recall of 83%, and an F1 score of 84%. This research addresses a gap in sentiment analysis studies related to refugee issues and provides valuable insights into public perceptions, which could inform policies and interventions aimed at improving refugee integration and support systems in Indonesia.

1. Introduction

Statelessness due to the Burma Citizenship Law enacted in 1982. This has forced many Rohingya to seek refuge in neighboring countries, with significant numbers arriving in Bangladesh and, more recently, in Aceh, Indonesia. The arrival of Rohingya refugees in Indonesia has elicited mixed reactions from the local populace, ranging from compassion to hostility [1]. Understanding these public sentiments is crucial for formulating effective responses to the refugee crisis.

Aceh has become the primary landing spot for these refugees. Although the distance between Aceh and Myanmar can be considered relatively close, it takes an average of about 23 days to reach Aceh's waters using traditional boats, during which time they are adrift at sea with minimal supplies [2]. The people of Aceh later called them "boat people" because of their long and arduous journey across the ocean to seek protection and help [3]. As the number of refugees arriving in Aceh increased, and with some refugees responding negatively to the assistance provided, negative sentiments began to arise among the Aceh community and Indonesians in general, leading to rejection actions against the Rohingya refugees [4], [5]. The National Human Rights Commission (Komnas HAM) of Indonesia conducted a series of monitoring processes regarding the presence of Rohingya refugees in Aceh from November to December 2023. The monitoring focused on refugee handling and the social dynamics, including rejection actions from parts of the community against the Rohingya refugees.

This study aims to fill a gap in the existing literature by applying sentiment analysis to assess public opinion on Rohingya refugees in Indonesia. Sentiment analysis is a process of automatically understanding and processing textual data to extract sentiment information contained in an opinion sentence [6]–[8]. One common machine learning method used for sentiment analysis is the Support Vector Machine (SVM) [9]–[11], as used in previous research by Ahmad Fatihin in 2022 on sentiment analysis of mobile application reviews [12]. In that study, sentiment analysis was conducted on the PeduliLindungi mobile application reviews given by users. Deep learning is also a machine learning method that can be used for sentiment analysis, often employing pre-trained models like Bidirectional Encoder Representations from Transformers (BERT). BERT is a pre-trained model trained using an English language corpus, while IndoBERT is trained using an Indonesian language corpus, as used in Bella Rahmatullah's 2021 research on sentiment analysis of the work-from-home implementation during the COVID-19 pandemic in Indonesia through Twitter opinions [13].

The specific objectives of this study are to (1) analyze the predominant public sentiments towards Rohingya refugees in Indonesia, (2) assess the performance of the IndoBERT model in sentiment classification, and (3) provide insights into the factors influencing these sentiments. By addressing these objectives, this research contributes to the fields of sentiment analysis and refugee studies, offering a deeper understanding of public attitudes that can inform policy-making and humanitarian efforts.

The significance of this study lies in its potential to influence policy and public opinion, providing a data-driven basis for interventions aimed at improving the lives of Rohingya refugees in Indonesia. By highlighting the sentiments of the Indonesian populace, this research can guide stakeholders, including policymakers, humanitarian organizations, and social workers, in their efforts to address the needs and challenges of the Rohingya community. This study focuses on sentiment data collected over a specific time frame from various online sources, ensuring a comprehensive analysis of public opinion. The geographical focus is limited to Indonesia, providing a detailed examination of sentiments within this context. The scope includes the application of the IndoBERT model, a state-of-the-art language model for the Indonesian language, ensuring the relevance and accuracy of the sentiment analysis.

The Rohingya refer to an ethnic group practicing Islam that has long lived in the Rakhine state. There are two versions of the origin of the term "Rohingya". One version states that it comes from "Rohang," which referred to Arakan (now Rakhine) in the past. Another version claims that the term "Rohingya" was coined by British researcher Francis Hamilton in the 18th century for all Muslim inhabitants in Arakan [14]. Rohingya used to be the second largest ethnic group inhabiting the state of Arakan. The state of Arakan has 5 districts and 17 cities, 3 of which, namely Buthidaung, Maungdaw, and Rathedaung are cities with the majority ethnic group being Rohingya. In the second world war, the Rohingyas supported England, this was in contrast to Myanmar which supported Japan [15].

The Rohingya people have been victims of genocide since 1978. The Myanmar government has committed human rights violations including mass killings, rape, forced seizure of Rohingya land and homes, torture of children, forced conversion to Buddhism, and destruction of ruins, marriage restrictions for other Rohingya, and many other human rights violations committed by Myanmar [16]. This is one of the contributions to the Rohingya's bad record for Myanmar. In 2015 things got worse, the Association for the Protection of Race and Religion (Ma Ba Tha) led by Buddhist monks succeeded in exploiting Myanmar's ultra-nationalist nature to persuade parliament to pass laws supporting Myanmar's Buddhist community [17].

Sentiment analysis is a multidisciplinary field involving psychology, sociology, Natural Language Processing (NLP), and machine learning [18]. Its primary function is to derive opinions expressed by the public on a topic to obtain insights on how to address that topic. In addition, sentiment analysis is carried out so that it can be used as a reference in improving a service, or improving product quality [19]. Based on the level, sentiment analysis is divided into 3, namely document level sentiment analysis, sentence level sentiment analysis, aspect-level sentiment analysis [20], [21].

The Indonesia Sentiment (InSet) lexicon is a lexicon-based method used to classify sentiments into three classes: positive, neutral, and negative. The InSet lexicon consists of 3609 positive words and 6609 negative words, with word scores ranging from -5 to +5 [22]. The InSet lexicon is also built to identify written opinions and categorize them into positive or negative opinions which can be used to analyze public sentiment towards a particular event topic, or product [23].

Bidirectional Encoder Representations from Transformers (BERT) is a pre-trained model introduced by Google in 2018. BERT is trained using the BookCorpus of 800 million words and the English Wikipedia of 2.500 million words [24]. BERT is based on the Transformer encoder, which uses an attention mechanism and an encoder-decoder framework but does not rely on Recurrent Neural Networks (RNN) [25].

IndoBERT is the Indonesian version of BERT. It is trained using more than 220 million Indonesian words from Indonesian Wikipedia (74 million words), several news articles (55 million words), and an Indonesian web corpus (90 million words) [26].

Confusion matrix is a matrix-shaped measurement tool used to obtain the amount of classification accuracy of the class with the algorithm used [27]. Basically, the confusion matrix contains information that compares the classification results performed by the system with the classification results that should be [28]. It provides values such as accuracy, precision, and recall. Accuracy is the ratio of correct prediction results [29]. Precision is the probability that a selected item is relevant and needed [30]. Recall is a value that shows the success rate or specificity to retrieve information correctly about negative class data or positive text content [31].

2. Method

The research methodology follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework show as Fig.1, encompassing the following stages:

- Business Understanding

  The first step involves defining the problem and objectives related to analyzing sentiments towards Rohingya refugees in Indonesia. This includes identifying the need to understand public sentiment to inform policy and humanitarian efforts, and setting specific goals for the sentiment analysis.

- Data Understanding

  In this phase, social media data relevant to the sentiments towards Rohingya refugees in Indonesia are collected. The sources include social media platforms and online news articles. The collected data are then explored to understand their characteristics, distribution, and relevance to the study. This step helps in identifying potential challenges and requirements for data preprocessing.

- Data Preparation

  This critical step involves preprocessing the collected data to make it suitable for analysis. The preprocessing techniques include:

  - Tokenizing : Splitting the text into individual tokens or words.
  - Cleaning : Removing noise such as special characters, punctuation, and irrelevant information.
  - Case Folding : Converting all text to lowercase to ensure uniformity.
  - Stemming : Reducing words to their root forms.
  - Filtering : Removing stop words and other non-informative words.
  - Labeling : Assigning sentiment labels to the text data using the Indonesia Sentiment (InSet) lexicon, which classifies sentiments into positive, neutral, or negative

- Modeling

  In this stage, the IndoBERT model is trained with the preprocessed data. The model is fine-tuned by adjusting hyperparameters such as batch size, learning rate, and the number of epochs to optimize performance. The modeling process involves.

  - Splitting the data into training and validation sets.
  - Initializing the IndoBERT model with pre-trained weights.

- Training the model on the training set while monitoring its performance on the validation set.
- Fine-tuning the hyperparameters to achieve the best possible performance

- Evaluation

  The trained model's performance is assessed using a confusion matrix to derive metrics such as accuracy, precision, recall, and F1-score. These metrics provide insights into the model's effectiveness in classifying sentiments correctly. The evaluation process involves:

  - Calculating the confusion matrix by comparing the predicted labels with the true labels.
  - Deriving accuracy, precision, recall, and F1-score from the confusion matrix to evaluate the model's performance comprehensively.
  - Interpreting these metrics to understand the strengths and weaknesses of the model

- Documentation and Reporting

  The final stage involves compiling the research findings and methodological details into a comprehensive report. This report includes:

  - An overview of the research objectives and problem statement.
  - Detailed descriptions of the data collection, preprocessing, and modeling steps.
  - Presentation of the model evaluation results, including tables and visualizations.
  - Discussions on the implications of the findings for policy-making and humanitarian efforts.
  - Recommendations for future research and potential applications of the study's findings

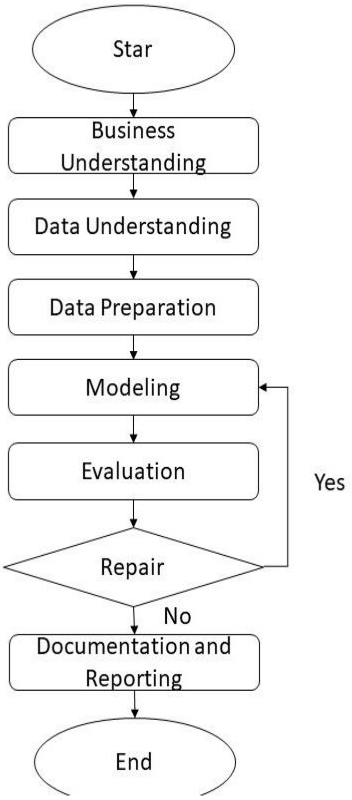

Fig. 1. CRISP DM Research Method

2.1. Business Understanding

Several things are done at this stage, such as understanding needs and goals from a business perspective, then interpreting knowledge in the form of defining problems in data mining and then determining plans and strategies to achieve data mining goals [32]. This stage involves understanding the problem object, which is the Rohingya refugees in Indonesia. The researchers also plan the research objectives and collect data based on the previous understanding of the problem object. The primary goal is to analyze public sentiments towards Rohingya refugees to inform policy-making and humanitarian efforts.

2.2. Data Understanding

Data understanding is conducted to familiarize with the data, identify data quality issues, find initial insights, and detect interesting subsets to form hypotheses about hidden information [33]. In this stage, researchers gain insights into the dataset obtained after data collection. This process, known as exploratory data analysis, involves analyzing word frequency in the sentiment data and creating visual representations of frequently occurring words. Additionally, irrelevant attributes in the dataset are removed.

2.3. Data Preparation

This stage involves preparing the dataset for modeling, including tokenizing, cleaning, case folding, stemming, filtering, and labeling [34]. The specific steps are:

- Tokenizing: Tokenizing is the stage of cutting words based on each word that makes it up [35]. Tokenization helps in searching for the frequency of information in a corpus.
- Cleaning: Removing or correcting words in the sentiment data that are misspelled or irrelevant, such as links, emoticons, etc.
- Case Folding: Converting all letters in the document to lowercase letters [36].
- Stemming: Stemming is the process of obtaining the root or base word of a word in a sentence by separating each word from its affixes, both prefixes and suffixes [37], e.g., "memulangkan" becomes "pulang".
- Filtering: Performing stopword removal and adding common abbreviations like "yg," "dg," "tdk," etc. Stopword removal aims to eliminate irrelevant words in a sentence based on a stopword list.
- Labeling: Assigning a class (positive, neutral, negative) to each sentiment in the dataset using the InSet Lexicon

2.4. Modeling

This phase involves the selection and application of various modeling techniques to obtain optimal results. After data preparation, the dataset is converted into an input format accepted by the IndoBERT model. Fine-tuning is performed to adapt the pre-trained IndoBERT model for sentiment classification using the given dataset. The optimizer used is Adam, with several hyperparameter trials to find the best values for:

- Batch size : 16 – 64
- Learning rate : 10-6 - 10-4
- Number of epochs : 5

2.5. Evaluation

Here, an analysis is conducted to determine whether the built model can meet business needs [38]. The classification results from the model are evaluated using a confusion matrix to assess the IndoBERT model's performance in sentiment analysis. Metrics such as accuracy, precision, recall, and F1-score are calculated to evaluate the model's effectiveness. The best hyperparameter set for the model is identified through this evaluation process.

2.6. Reporting and Documentation

The research process and results are documented and compiled into a final report. This report provides clear explanations and detailed descriptions of the methodologies, findings, and implications for those interested in the sentiment analysis results of Rohingya refugees in Indonesia. The report includes visual representations, tables, and comprehensive discussions on the significance of the findings.

3. Results and Discussion

3.1. Data Condition

Out of 1258 collected entries, 1081 (86%) were used for analysis after excluding entries with irrelevant or no sentiments. The excluded data often consisted of only emoticons or responses not directly related to the Rohingya issue. Examples of excluded data include entries that contained solely emojis or comments that were general and not specific to the Rohingya context.

- Sentiment Distribution

  A word cloud analysis revealed frequent negative words like 'usir' (expel) and 'tolak' (reject), while positive words included 'bantu' (help) and 'kasih' (give). The labeled dataset showed that 62.1% of sentiments were negative, 20.9% positive, and 17% neutral. This distribution indicates a predominantly negative sentiment towards the Rohingya refugees among the Indonesian populace. The visual representation of this data through word clouds highlights the prominence of specific negative and positive terms, providing a clearer picture of public opinion.

- Model Performance

  The IndoBERT model, using hyperparameters of batch size 64 and learning rate 5×10-5, achieved the highest accuracy of 100% for training, 71% for validation, and 74% for testing. Precision, recall, and F1-score were calculated to provide a comprehensive evaluation of the model's performance. The results indicated better performance for positive and negative classes compared to the neutral class. Specifically, the precision, recall, and F1-score for the positive class were 76%, 73%, and 74%, respectively, while for the negative class, these metrics were 78%, 75%, and 76%. The neutral class showed lower performance with precision, recall, and F1-score values of 68%, 65%, and 66%, respectively.

Based on the obtained data, a total of 1,258 entries were collected, of which 1,081 (86%) were used. This is because some data entries did not display any sentiment. Sentiments that were removed often contained only emoticons, which are not relevant to be considered as sentiments. Additionally, some removed sentiments were not related to the issue at hand, as they were responses to other people's sentiments. Table. 1 are examples of the data that were excluded.

Table.1 Sentiment deleted

Number	UserName	Sentiment
1	namucul	🗿
2	miwa_rd03	@miwa_rd03 canda yatim✌🙋‍♀️
3	titi_mukti_	👍👍👍👍👍👍👍

Entries number 1 and 3 are examples of data that contain only emoticons, making it difficult to label them as positive, neutral, or negative sentiments. In contrast, entry number 2 is an example of a sentiment expressed in response to someone else's sentiment. Since this sentiment does not directly relate to the issue at hand, it was excluded.

3.2. Data Understanding

In the understanding phase, comprehension of the obtained dataset was carried out. The dataset contained seven attributes, including id as a unique identifier for each sentiment, likesCount representing the number of likes on the given comment, ownerProfilePicUrl as the attribute link to the user's profile picture, ownerUsername representing the user's account name, postUrl as the link address of the post where the comment was made, text as the content of the comment serving as the user's sentiment, and timestamp as the time the comment was sent. Some attributes were removed as they were deemed irrelevant for the sentiment analysis process, namely id, likesCount, ownerProfilePicUrl, ownerUsername, postUrl, and timestamp.

Additionally, a wordcloud was created to identify the most frequently occurring words in the sentiment data related to the Rohingya issue in Indonesia. As shown in Fig. 2, the words with negative connotations that appeared frequently were "usir" (expel) and "tolak" (reject). In contrast, the words with positive connotations that appeared frequently were "bantu" (help) and "kasih" (give).

Fig. 2. Wordcloud

3.3. Data Preparation Result

During the data preparation process, several steps were undertaken to prepare the dataset for the modeling phase. These steps included case folding, tokenizing, normalization, stemming, and filtering using stopwords removal and common abbreviations. Table. 2 are some results from the data preparation process.

Table.2 Step of pre-processing

Text	Case Folding	Tokenizing	Normalization	Stemming	Filtering
Harus dibantu	harus dibantu	['harus', 'dibantu']	['harus', 'dibantu']	['harus', 'bantu']	['bantu']
Naikkan aja ke kapal pulangkan tu ke asalnya. Nyusahin aja	naikkan aja ke kapal pulangkan tu ke asalnya nyusahin aja	['naikkan', 'aja', 'ke', 'kapal', 'pulangkan', 'tu', 'ke', 'asalnya', 'nyusahin', 'aja']	['naikkan', 'saja', 'ke', 'kapal', 'pulangkan', 'itu', 'ke', 'asalnya', 'menyusahkan', 'saja']	['naik', 'saja', 'ke', 'kapal', 'pulang', 'itu', 'ke', 'asal', 'susah', 'saja']	['kapal', 'pulang', 'susah']
Belum apa sudah nyusahin. Kalau mereka mogok makan biarin gak usah dikasih makan biar mati sekalian	belum apa sudah nyusahin kalau mereka mogok makan biarin gak usah dikasih makan biar mati sekalian	['belum', 'apa', 'sudah', 'nyusahin', 'kalau', 'mereka', 'mogok', 'makan', 'biarin', 'gak', 'usah', 'dikasih', 'makan', 'biar', 'mati', 'sekalian']	['belum', 'apa', 'sudah', 'menyusahkan', 'kalau', 'mereka', 'mogok', 'makan', 'biarkan', 'enggak', 'usah', 'dikasih', 'makan', 'biar', 'mati', 'sekalian']	['belum', 'apa', 'sudah', 'susah', 'kalau', 'mereka', 'mogok', 'makan', 'biar', 'enggak', 'usah', 'kasih', 'makan', 'biar', 'mati', 'sekali']	['susah', 'mogok', 'makan', 'kasih', 'makan', 'mati']

After preprocessing the sentiments, labeling was performed using the Inset lexicon. Table. 3 are some results from the labeling process.

Table.3 Labelling result

Sentiment	Polarity Score	Label
Ya jelas beda lah, di Palestina jelas-jelas ada perang membabi buta, sedang rohingya sudah aman di bangladesh tidak ada pembantaian malah kabur pindah negara. Bayar belasan bahkan puluha juta, itu bukan darurat, tapi sindikat!	-24	Negative
Dulu masyarakat tidak tahu ternyata ada udang dibalik batu. masyarakat tahunya menolong orang yang membutuhkan. ternyata kebaikan mereka dimanfaatkan oleh pengacau dunia.	0	Neutral
setuju bapak mending dibuatkan rumah pengungsi di hambalang	1	Positive
Katanya ikhlas kok setengah-setengah mau jadi presiden kok tidak ada kemanusiaan yang adil dan beradab.	14	Positive

The sentiment labeling results were reevaluated, revealing several instances deemed imprecise by the researchers, as exemplified show as Table. 4:

Table.4 Mislabeling

Sentiment	Polarity Score	Label
Suruh mereka angkat kaki dari Pulau Sumatera	1	Positive
Kalo Indonesia nggak ikut PBB beda cerita bro	-4	Negative
Kerja juga gak, makan juga harap dari pemberian, uang juga dikasih, enak bener, kalian bukan tanggung jawab bangsa indonesia ya	2	Positive

As observed in Table 4, the researchers found the first and third sentiments to lean more towards negative sentiment, whereas the second sentiment was perceived to incline towards neutrality.

The polarity score is a value assigned to a sentiment that determines whether the statement falls under positive, neutral, or negative labels. It is categorized as positive sentiment if the polarity score is greater than 0, neutral if it equals 0, and negative if it is less than 0. The distribution of sentiment labels show as Fig. 3.

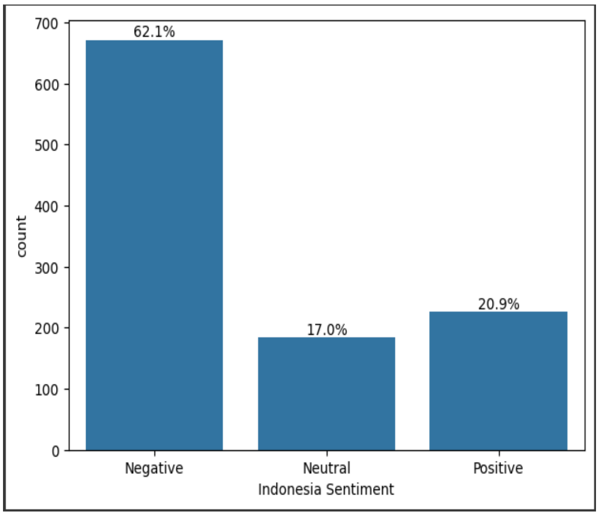

Fig. 3. The distribution of sentiment labels

After labeling the entire dataset, it was found that the number of sentiments labeled as negative was significantly higher, comprising 62.1% or 671 out of 1081 data. This was followed by positive labels, accounting for 20.9% or 226 out of 1081 data. Finally, neutral labels constituted 17% or 184 out of 1081 data.

3.4. Sentiment Classification Results

In this study, the sentiment classification results will be categorized based on the hyperparameters employed. The model utilized in this research is IndoBERT (Indonesian Bidirectional Encoder Representations from Transformers). Based on several studies, the learning rate used is [3×10^{-6}, 5×10^{-5}, 10^{-4}], for batch size 16 and 64, and epochs 5. Evaluation show as Table. 5.

Table.5 Evaluation

	Training Accuracy	*Validation Accuracy*	*Testing Accuracy*
Batch size 64 , lr 5×10^{-5}	1	0.71	0.74
Batch size 16 , lr 5×10^{-5}	0.93	0.6	0.61
Batch size 64 , lr 3×10^{-6}	0.63	0.64	0.61
Batch size 16 , lr 3×10^{-6}	0.83	0.66	0.72
Batch size 64 , lr 10^{-4}	0.96	0.68	0.63
Batch size 16 , lr 10^{-4}	0.9	0.67	0.65

From the model evaluation results obtained by comparing several hyperparameters using the IndoBERT model on the dataset of Indonesian public sentiment towards Rohingya, it was found that the hyperparameter with the highest accuracy value was batch size 64 and learning rate 5×10^{-5} with training, validation, and testing accuracies sequentially as follows 1, 0.71, dan 0.74. The model performance show as Fig. 4.

	precision	recall	f1-score	support
negative	0.78	0.88	0.83	33
neutral	0.40	0.33	0.36	6
positive	0.75	0.60	0.67	15
accuracy			0.74	54
macro avg	0.64	0.60	0.62	54
weighted avg	0.73	0.74	0.73	54

Fig. 4. The model performance with batch size 64 and learning rate 5×10^{-5} on testing

For the parameter values related to precision, recall, and F1-score, the values obtained using the hyperparameters batch size 64 and learning rate 5×10^{-5} yielded better values compared to other tested hyperparameters. However, the parameter values obtained for the neutral class are relatively low compared to other classes. Nonetheless, these results are still superior compared to the parameter values for the neutral class obtained from testing other hyperparameters below batch size 16 and learning rate 3×10^{-6} with precision, recall, and F1-score values in sequence for the neutral class being 0.43, 0.5, and 0.46.

4. Conclusion

The study concludes that public sentiments in Indonesia towards Rohingya refugees predominantly lean towards negativity. The sentiment analysis using the IndoBERT model provides valuable insights into public opinion, which can inform policymakers and humanitarian organizations in addressing the refugee crisis. Future research should consider expanding the dataset, ensuring balanced sentiment labels, and exploring alternative labeling systems for improved accuracy. 1) Sentiments prevalent among the Indonesian populace from November 2023 to January 30, 2024, regarding Rohingya refugees arriving in Indonesia tend to lean towards negativity. This negative sentiment trend indicates a significant portion of Indonesians rejecting the presence of Rohingya refugees in their homeland; 2) In sentiment analysis modeling using IndoBERT, the selection of hyperparameters plays a significant role in the model's outcomes. This is evident through the confusion matrix results generated for different sets of hyperparameter values; 3) The performance of the IndoBERT model in analyzing sentiment among the Indonesian populace regarding Rohingya refugees, using a data split of 70% for training, 25% for validation, and 5% for testing, yielded accuracies of 100% for training, 71% for validation, and 74% for testing. These results were achieved using the hyperparameters for epochs set at 5, batch size set at 64, and learning rate set at 5×10-5. Suggestions for future research include; 1) Ensure an equal number of samples for positive, neutral, and negative sentiments to improve model performance and reliability. This could help address potential biases in the dataset and provide a more accurate representation of public sentiment; 2) Conduct comparative studies to evaluate the performance of IndoBERT against other models, including different variations of BERT models. This would provide valuable insights into the strengths and weaknesses of each approach and identify the most effective model for sentiment analysis in this context; 3) Explore different lexicons or develop custom labeling methodologies tailored to the specific context of the sentiment analysis task. Considering discrepancies observed in sentiment labeling, experimenting with various lexicons or creating a bespoke labeling system could enhance the accuracy and relevance of the analysis; 4) Future research should consider expanding the dataset to include more diverse sources of sentiment data. Incorporating data from various social media platforms, news outlets, and other relevant sources can provide a more comprehensive understanding of public opinion; 5) Conducting a temporal analysis of sentiment trends over an extended period could provide insights into how public sentiments towards Rohingya refugees evolve over time. This would help in understanding the impact of specific events or policies on public opinion; 6) Further research could explore the policy implications of the findings, providing actionable recommendations for policymakers and humanitarian organizations. Understanding the root causes of negative sentiments and identifying strategies to address them could help in fostering a more supportive environment for refugees; 7) Investigate effective engagement strategies to improve public perception and support for Rohingya refugees. This could include public awareness campaigns, educational programs, and community engagement initiatives aimed at promoting social cohesion and inclusivity.

References

[1] A. Hakim and Arini Nur Annisa, "Legal Governance Of Rohingya Refugee Affairs In Aceh: A Constitutional Perspective," *J. Sultan Ris. Huk. Tata Negara*, vol. 2, no. 2, pp. 21–32, Feb. 2024, doi: 10.35905/sultan_htn.v2i2.7990.

[2] M. A. Hamdi, H. Maulidia, and H. Firlana, "Phenomenon Of Asylum Seekers And Refugees Of Rohingya Ethnic In Indonesia," *J. Ilm. Kaji. Keimigrasian*, vol. 6, no. 1, pp. 55–71, Aug. 2023, doi: 10.52617/jikk.v6i1.404.

[3] A. Azis, "The Refugees Vanish: Rohingya Movement, Emergency's Temporality and Violence of the Indonesian Humanitarian Border," in *IMISCOE Research Series*, Springer, Cham, 2023, pp. 177–193, doi: 10.1007/978-3-031-25748-3_10.

[4] M. Riza, F. Anas, A. Sukendro, I. Malik, P. Widodo, and A. A. Chandra, "Analysis of Dynamic Framework to Prevent Potential Post-Migration Conflict of Rohingya Ethnic in Aceh," *J. Law, Polit. Humanit.*, vol. 4, no. 6, pp. 1978–1989, Sep. 2024. [Online]. Available at: https://dinastires.org/JLPH/article/view/572.

[5] H. Durohman, Fajar Andrian Sutisna, and Muhamad Ridho, "The Analysis of Pros and Cons Regarding the Arrival of Rohingya Refugees: Religious, Humanitarian, and Socio-Economic Implication," *J. Islam. Civiliz.*, vol. 5, no. 2, pp. 237–259, May 2024, doi: 10.33086/jic.v5i2.5788.

[6] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif. Intell. Rev. 2022 557*, vol. 55, no. 7, pp. 5731–5780, Feb. 2022, doi: 10.1007/S10462-022-10144-1.

[7] A. A. Q. Aqlan, B. Manjula, and R. Lakshman Naik, "A Study of Sentiment Analysis: Concepts, Techniques, and Challenges," in *Lecture Notes on Data Engineering and Communications Technologies*, vol. 28, Springer, Singapore, 2019, pp. 147–162, doi: 10.1007/978-981-13-6459-4_16.

[8] D. Gritsenko, M. Wijermars, and M. Kopotev, *The Palgrave Handbook of Digital Russia Studies*. 2020, doi: 10.1007/978-3-030-42855-6.

[9] W. Bourequat and H. Mourad, "Sentiment Analysis Approach for Analyzing iPhone Release using Support Vector Machine," *Int. J. Adv. Data Inf. Syst.*, vol. 2, no. 1, pp. 36–44, Apr. 2021, doi: 10.25008/ijadis.v2i1.1216.

[10] A. Gupta, P. Tyagi, T. Choudhury, and M. Shamoon, "Sentiment Analysis Using Support Vector Machine," in *2019 International Conference on contemporary Computing and Informatics (IC3I)*, Dec. 2019, pp. 49–53, doi: 10.1109/IC3I46837.2019.9055645.

[11] K.-X. Han, W. Chien, C.-C. Chiu, and Y.-T. Cheng, "Application of Support Vector Machine (SVM) in the Sentiment Analysis of Twitter DataSet," *Appl. Sci.*, vol. 10, no. 3, p. 1125, Feb. 2020, doi: 10.3390/app10031125.

[12] D. Affianty, "The Impact Of The Rohingya Issue On Violent Extremism In Indonesia," Universitas Muhammadiyah Jakarta, p. 23, 2019. [Online]. Available at: https://repository.umj.ac.id/8556/1/Laporan Penelitian-Dampak Isu Rohingya Terhadap Ekstremisme Kekerasan di Indonesia.pdf.

[13] B. Rahmatullah, "Sentiment Analysis of Work From Home Implementation in Indonesia During the COVID-19 Pandemic Using IndoBERT,", pp. 1-101, 2021. [Online]. Available at: https://repository.its.ac.id/91493/1/06111640000078-Undergraduate_Thesis.pdf.

[14] J. Nababan, "The Role of the International Criminal Court (ICC) in Handling Crimes Against Humanity: A Case Study of the Rohingya in Myanmar (2016)," *J. Univ. Komput. Indones.*, vol. 2, no. 2, pp. 1–13, 2018. [Online]. Available at: https://elibrary.unikom.ac.id/id/eprint/1905/13/unikom.

[15] A. Roy Chowdhury, "An 'un-imagined community': the entangled genealogy of an exclusivist nationalism in Myanmar and the Rohingya refugee crisis," *Soc. Identities*, vol. 26, no. 5, pp. 590–607, Sep. 2020, doi: 10.1080/13504630.2020.1782731.

[16] V. Septia Renanda, D. Natasyafira, A. Julia Kusuma, Z. Delia Reviska, and M. Putri Winarti, "Legal Protection For The Rohingya People From The Perspective Of Human Rights And International Law," *Sibatik J. J. Ilm. Bid. Sos. Ekon. Budaya, Teknol. dan Pendidik.*, vol. 2, no. 1, pp. 143–152, Dec. 2022, doi: 10.54443/sibatik.v2i1.510.

[17] H. Wahyudi and N. Niko, "Indonesian Diplomacy as a Muslim Majority Democracy in Responding to Human Rights Violations in Myanmar," *POLITEA*, vol. 6, no. 1, p. 107, Jun. 2023, doi: 10.21043/politea.v6i1.20350.

[18] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: a tertiary study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, Oct. 2021, doi: 10.1007/s10462-021-09973-3.

[19] P. K. Jain, R. Pamula, and G. Srivastava, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Comput. Sci. Rev.*, vol. 41, p. 100413, Aug. 2021, doi: 10.1016/j.cosrev.2021.100413.

[20] M. Liu, F. Zhou, K. Chen, and Y. Zhao, "Co-attention networks based on aspect and context for aspect-level sentiment analysis," *Knowledge-Based Syst.*, vol. 217, p. 106810, Apr. 2021, doi: 10.1016/j.knosys.2021.106810.

[21] Y. Lv *et al.*, "Aspect-level sentiment analysis using context and aspect memory network," *Neurocomputing*, vol. 428, pp. 195–205, Mar. 2021, doi: 10.1016/j.neucom.2020.11.049.

[22] F. Koto and G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," in *2017 International Conference on Asian Language Processing (IALP)*, Dec. 2017, vol. 2018-Janua, pp. 391–394, doi: 10.1109/IALP.2017.8300625.

[23] C. Zucco, B. Calabrese, G. Agapito, P. H. Guzzi, and M. Cannataro, "Sentiment analysis for mining texts and social networks data: Methods and tools," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 1, p. e1333, Jan. 2020, doi: 10.1002/WIDM.1333.

[24] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North*, 2019, pp. 4171–4186, doi: 10.18653/v1/N19-1423.

[25] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer models for text-based emotion detection: a review of BERT-based approaches," *Artif. Intell. Rev.*, vol. 54, no. 8, pp. 5789–5829, Dec. 2021, doi: 10.1007/s10462-021-09958-2.

[26] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics*, 2020, pp. 757–770, doi: 10.18653/v1/2020.coling-main.66.

[27] W. A. Hamwi and M. M. Almustafa, "Development and integration of VGG and dense transfer-learning systems supported with diverse lung images for discovery of the Coronavirus identity," *Informatics Med. Unlocked*, vol. 32, p. 101004, Jan. 2022, doi: 10.1016/j.imu.2022.101004.

[28] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Comput. Oper. Res.*, vol. 152, p. 106131, Apr. 2023, doi: 10.1016/j.cor.2022.106131.

[29] C. Tofallis, "A better measure of relative prediction accuracy for model selection and model estimation," *J. Oper. Res. Soc.*, vol. 66, no. 8, pp. 1352–1362, 2015, doi: 10.1057/jors.2014.103.

[30] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," pp. 37–63, 2020, [Online]. Available at: http://arxiv.org/abs/2010.16061.

[31] B. Satya, M. H. S J, M. Rahardi, and F. F. Abdulloh, "Sentiment Analysis of Review Sestyc Using Support Vector Machine, Naive Bayes, and Logistic Regression Algorithm," in *2022 5th International Conference on Information and Communications Technology (ICOIACT)*, Aug. 2022, pp. 188–193, doi: 10.1109/ICOIACT55506.2022.9972046.

[32] T. Silwattananusarn, “Data Mining and Its Applications for Knowledge Management : A Literature Review from 2007 to 2012,” *Int. J. Data Min. Knowl. Manag. Process*, vol. 2, no. 5, pp. 13–24, Sep. 2012, doi: 10.5121/ijdkp.2012.2502.

[33] Lathifah, “Cross-Industry Standard Process For Data Mining (Crisp-Dm) To Find Association Patterns In Data Tracer Study Of Islamic College Graduates (Case Study: Uin Syarif Hidayatullah Jakarta) Cross-Industry Standard Process For Data Mining (Crisp-Dm) To Find Asso,” Uin Syarif Hidayatullah Jakarta, p.223, 2023, doi: 10.1109/CITSM60085.2023.10455691.

[34] P. P. W. Suyitno, R. E. Indrajit, and M. Fauzi, “Application Of Data Mining In Handling Congestion In Jakarta,” *Ikra-Ith Inform. J. Komput. dan Inform.*, vol. 1, no. 2, pp. 53–60, 2017. [Online]. Available at: http://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/110.

[35] A. Rai and S. Borah, “Study of Various Methods for Tokenization,” in *Lecture Notes in Networks and Systems*, vol. 137, Springer, Singapore, 2021, pp. 193–200, doi: 10.1007/978-981-15-6198-6_18.

[36] J. A. Pamungkas, Y. Priyadi, and M. J. Alibasa, “Measurement of Similarity Between Requirement Elicitation and Requirement Specification Using Text Pre-Processing in the Cinemaloka Application,” in *2022 IEEE World AI IoT Congress (AIIoT)*, Jun. 2022, pp. 672–678, doi: 10.1109/AIIoT54504.2022.9817193.

[37] A. Jabbar, S. Iqbal, M. I. Tamimy, A. Rehman, S. A. Bahaj, and T. Saba, “An Analytical Analysis of Text Stemming Methodologies in Information Retrieval and Natural Language Processing Systems,” *IEEE Access*, vol. 11, pp. 133681–133702, 2023, doi: 10.1109/ACCESS.2023.3332710.

[38] V. Hananto, “Analysis of Determining Data Mining Methods for Predicting Student Graduation as a Support for Educational Efficiency Figures,” *J. Ilm. SCROLL*, vol. 5, no. 1, pp. 1–11, 2017, [Online]. Available at: http://repository.dinamika.ac.id/id/eprint/2122/4/.