Sentiment Analysis of Indonesian YouTube Reviews About Lesbian, Gay, Bisexual, and Transgender (LGBT) using IndoBERT Fine Tuning

Teddy Oswari[a1], Murniyati [a2], Trityanti Yusnitasari[a3], Nurasiah[a4], Seviyanti Wijay[a5]

[a]Information System, Gunadarma University
Jl. Margonda Raya 100 Depok, Indonesia
[1]teddyoswari@yahoo.com
[2]murnirk77@gmail.com (Corresponding author)
[3]tyusnita@staff.gunadarma.ac.id, [4]nurasiah2208@gmail.com
[5]vika@gmail.com

Abstract

Lesbian, gay, Bisexual, and Transgender (LGBT) is an individual who has a sexual orientation or gender identity that is different from the heterosexual majority. The LGBT community now dares to appear openly on social media; nowadays, social media is used as a source of information and a place to provide comments. The Indonesian state generally still views the LGBT community as deviant behavior. This research was conducted to understand Indonesian society's views on LGBT through YouTube and social media. The text mining method analyzes and classifies the counter or pro sentences expressed in the comments. The model used in this research is IndoBERT because the research object studied is Indonesian. IndoBERT is part of the Bidirectional Encoder Representation From Transformers (BERT) model. The data sources used were 1,493 data. The stages carried out in this research included the preprocessing stage, which included case folding, data cleaning, tokenization, stopword removal, stemming, and normalization, then the data labeling stage, and finally, the model building stage with IndoBERT Fine Tuning. The level of accuracy achieved using IndoBERT is 74%.

Keywords: LGBT, IndoBERT, Sentiment Analysis, Youtube, BERT

1. Introduction

The growth of LGBT people in Indonesia has attracted the attention of many parties and has become a hot topic of conversation in society, primarily through social media. Lesbian, gay, Bisexual, and Transgender (LGBT) has become an issue that attracts public attention because they are increasingly daring to appear openly, primarily through promotional activities and advertisements spread on social media. On social media, the LGBT movement has generated various comments, both supportive and opposing [1]. In Indonesia, in general, LGBT is still seen by some people as deviant behavior because it is considered to deviate from conventional sexual orientation. The debate regarding LGBT covers different points of view; where some people hope that LGBT existence is respected as part of human rights, while others consider this behavior to be deviant and sinful [2].

Social media, especially the YouTube platform, is the primary source of information and where many public comments and opinions are expressed [3]. The comments column on YouTube is a forum for users to respond to uploaded content. Sentiment analysis of comments on YouTube can be an alternative for understanding public opinion and sentiment towards content [4].

Research focusing on Sentiment Analysis of Cyberbullying on Twitter produced an accuracy of 81% with an f1 score of 80.67% [5]. The study examines the object of Sentiment Analysis of English-language Film Reviews using the Bidirectional Encoder Representations from Transformers Approach. Dataset collection was carried out by downloading data from the Cornelledu website, where the data used was the Movie Review Polarity Dataset v2.0 data. There are 2000 film reviews, of which 1000 are positive, and the remaining 1000 data are negative reviews. BERT-base is proven to be used not only for short sentences to be classified. Even

though only the final 128 characters were taken from the total characters in the document, BERT-base could still carry out training well so that the accuracy results were relatively high, namely 73.7% [6] [7].

Several studies have researched sentiment analysis or the implementation of sentiment analysis using various methods. The BERT algorithm has slightly better classification performance than the random forest algorithm; simulation results show that the random forest algorithm has significantly lower computing time compared to the BERT algorithm in the sentiment analysis of Twitter users [8]; the following is some previous research related to the study conducted sentiment analysis on Twitter social media using Naïve Bayes. The method used is this method with 1,579 data taken from Twitter. The accuracy level is 95%, with a positive polarity value of 77% and a negative polarity of 150.2%. To confirm the results of this research, a K-fold cross-validation test was also carried out with k of 10, which resulted in an accuracy value of 87.58% [9].

Research was conducted on implementing the SVM algorithm for sentiment analysis regarding LGBT reviews. The method used in this research is the support vector machine method. The classification process uses three kernels: linear, polynomial, and Radical Basic Function (RBF). The results obtained by the SVM accuracy process produced 74% accuracy in the linear kernel with experimental data of 90%: 10% and 74% and in the RBF kernel with C-100 gamma=0.01 [10].

Research was conducted regarding sentiment analysis of YouTube video comments using support vector machines. In this research, the support vector machine method was used, and classification was carried out on Indonesian language YouTube comments with the themes of daily vlogs, culinary videos, unboxing, hack/DIY videos, and music cover videos. The results that have been carried out using the support vector machine algorithm, which uses tf-idftf-IDF feature extraction and also uses a polynomial kernel, have an accuracy of 86%, the precision of 87%, recall of 99%, and f1-score of 100% [11].

Other research related to sentiment analysis regarding the increase in prices of necessities on YouTube and social media using the SVM algorithm. This research uses a support vector machine to classify positive and negative comment labels and uses a technique for balancing the number of labels in the data, namely SMOTE. During the testing of the four SVM classification models, linear SVM with SMOTE achieved the highest accuracy with 86.33%, precision 75%, recall 66.67%, and f1-score 70.59%. Testing was also carried out on validation data (new test data outside of the dataset), resulting in an accuracy value of 72% [12].

The Sentiment Analysis of Electronic System Operator (PSE) Policies research used the Bidirectional Encoder Representations from Transformers (BERT) Algorithm. This research used a dataset of 5,016 and used hyperparameters. The batch size was 16, and the epoch was 5, resulting in accuracy levels of 69%, 55%, and 55% [13]. One way to manage hate speech on social media is to classify a sentence to determine whether the sentence is hate speech or an average sentence. The model created for classification uses the feedforward neural network method with IndoBERT. Based on the test results, the model developed using the feedforward neural network method with IndoBERT provides the best accuracy of 89.52% [14]

The novelty of this research lies in producing a sentiment analysis of Indonesian YouTube reviews regarding LGBT using IndoBert. Thus, this study is expected to contribute to developing sentiment analysis classification modeling using IndoBert.

2. Research Methods

The CRISP-DM (Cross-Industry Standard Process for Data Mining) model provides an overview of the lifecycle of a data mining project. CRISP-DM consists of 6 stages: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment, as shown in Figure 1. The research steps using the CRISP-DM Method are shown in Figure 1.
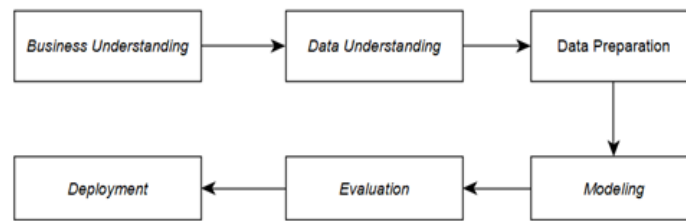
Figure 1. Research flow using the CRISP-DM method

In the first stage, business understanding in CRISP-DM aims to understand the purpose of creating the program that will be used. In the second stage, data understanding is used to analyze the data using a dataset from Kaggle.com [15]. In the third stage, the aim is to get good data so that it will produce good results, too [16]. Data preparation is carried out to clean and prepare the data. In the fourth stage, there is a modeling stage where the data will be processed using the algorithm BERT, which will go through a fine-tuning process by adjusting hyperparameters when the training stage begins. The fine-tuning process is done to find a configuration that works on the dataset [17][18]. That will be used, and in the fifth stage, evaluation is a model evaluation to determine the accuracy of the results of the model used. In the final stage of deployment, the model used will be implemented. This method is used in classification models to visualize and calculate the number of correct and incorrect predictions made by the model [19]. CRISP-DM-based model for data mining in the process of predicting assembly cycle time. The developed solution has been evaluated using accuracy. Industrial data will be a part of the methodology that allows the assembly estimate. Time of a finished product at the quotation stage [20].

3. Result and Discussion

3.1. Business Understanding

At the business understanding stage of this program, the aim is to classify sentences related to LGBT issues so that the public can determine public opinion and analyze sentences classified as being in the contra (positive) or pro (negative) category toward LGBT.

3.2. Data Understanding

In the data understanding stage, the research used a dataset from Kaggle.com, accessed on May 30, 2023. This dataset contains YouTube comments about LGBT. It has two attributes: "comment," which includes comments, and "sentiment," which is the label of the comment sentence. A total of 1,493 data points were used, with a data split of 80% for training data, 10 % for validation, and 10% for testing data, as shown in Figure 2.

Figure 2. Attributes in the Dataset

A word cloud can be created using the data shown in Figure 3 below.

Figure 3. Wordcloud

3.3. Data Preparation

The aim of the data preparation stage is to obtain good data so that it will produce good results. To do this, the preprocessing stage is carried out. The case folding, data cleansing, tokenizing, stopword, and stemming processes will be carried out at this stage. The preprocessing stage is shown in Figure 4 below.

Figure 4. Preprocessing Stage

3.3.1. Case Folding

This stage is the process of changing uppercase letters to lowercase. Initial sentence to case folding result as shown in Table 1:

Table 1. Case Folding

Initial Sentence	Case Folding Results
Toleransi?? Oke, Kami menghormati pilihan hidup anda Tapi untuk mendukung LGBT?? Tentu saja tidak	toleransi?? oke, kami menghormati pilihan hidup anda tapi untuk mendukung lgbt?? tentu saja tidak
Tuhan menciptakan Adan dan Hawa, bukan Adam dan Ipul ☺	tuhan menciptakan adam dan hawa, bukan adam dan ipul ☺

Channel orang2 sakit/ menyimpang	channel orang2 sakit/ menyimpang

3.3.2.   Data Cleansing

The data cleansing stage removes attributes in sentences, such as emojis, hashtags, numbers, punctuation marks, and words that do not provide information. Table 2 shows the case folding result to the data cleaning result. This part presents results or findings.

Table 2. Data Cleansing

Case Folding Results	Data Cleaning Results
toleransi?? oke, kami menghormati pilihan hidup anda tapi untuk mendukung lgbt?? tentu saja tidak	toleransi oke kami menghormati pilihan hidup anda tapi untuk mendukung lgbt tentu saja tidak
tuhan menciptakan adam dan hawa, bukan adam dan ipul ☺	tuhan menciptakan adam dan hawa, bukan adam dan ipul
channel orang2 sakit/ menyimpang	channel orang sakit menyimpang

In the table, the results of the case folding process will be used to process data cleansing. This process will remove attributes in sentences and delete emojis, hashtags, numbers, punctuation marks, and excessive spaces in sentences; the example sentences in the table above remove punctuation marks and emojis.

3.3.3.   Tokenizing

At this tokenizing stage, the text is divided into sentences and word-by-word units. Data cleaning results in tokenizing, as shown in Table 3.

Table 3. Tokenizing

Data Cleaning Results	Tokenizing Results
toleransi oke kami menghormati pilihan hidup anda tapi untuk mendukung lgbt tentu saja tidak	['toleransi', 'oke', 'kami', 'menghormati,' 'pilihan,' 'hidup', 'anda', 'tapi', 'untuk', 'mendukung', 'lgbt', 'tentu', 'saja', 'tidak']
tuhan menciptakan adam dan hawa, bukan adam dan ipul	['tuhan', 'menciptakan', 'adam', 'dan', 'hawa', 'bukan', 'adam', 'dan', 'ipul']
channel orang sakit menyimpang	['channel', 'orang', 'sakit', 'menyimpang']

The table's data cleansing results will be processed to the tokenizing stage by breaking sentences into word units.

3.3.4.   Stopword and Stemming

The stopword stage will delete words that are not needed, and the stemming stage will change the words into their basic form. This stemming process uses an Indonesian literary library—the tokenizing results to stopword and Stemming results as shown in Table 4.

Table 4. Stopword and Stemming

Tokenizing Results	Stopword and Stemming Results
['toleransi', 'oke', 'kami', 'menghormati,' 'pilihan,' 'hidup', 'anda', 'tapi', 'untuk', 'mendukung', 'lgbt', 'tentu', 'saja', 'tidak']	toleransi oke hormat pilih hidup dukung lgbt

['tuhan', 'menciptakan', 'adam', 'dan', 'hawa', 'bukan', 'adam', 'dan', 'ipul']	tuhan cipta adam hawa adam ipul
['channel', 'orang', 'sakit', 'menyimpang']	channel orang sakit simpang

In the table above, the results of tokenizing will carry out a combination of stopword and stemming processes. The table above will remove unnecessary words and change them into their basic form.

3.3.5. Normalization

In the normalization stage, the dataset containing non-standard words is converted into standard words or according to spelling. This is done because quite a lot of sentences use non-standard words such as: "enggak", "ngga", and "ga". If the word does not go through the normalization stage, then the system will consider the words "ga", "enggak", "ngga", and "gak" are different words, even though these words have the same meaning, namely "tidak". as showed in Figure 5.
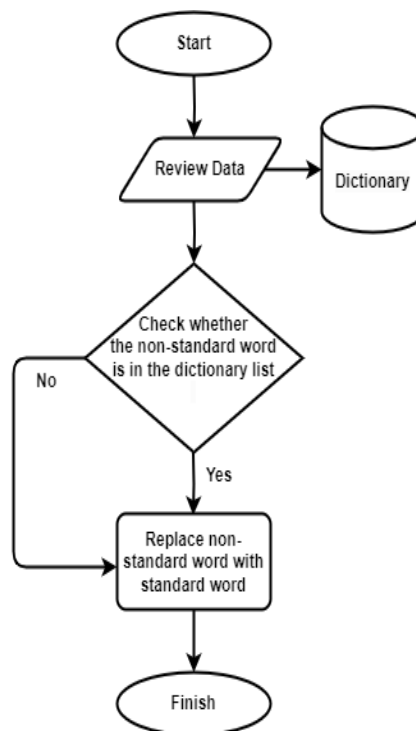

Figure 5. Normalization

3.4. Lexicon Based Labeling

Lexicon Based Labeling is a method of sentiment analysis where the process of determining sentiment is carried out through an approach where the words in the text are analyzed based on the sentiment value associated with those words. A lexicon is a reference to determine whether a word or sentence has a positive, neutral, or negative sentiment. If a text has a polarity value >0, then it is classified as text with a positive sentiment; if a text has a polarity value of =0, then it is classified as a text with a neutral sentiment; and if a text has a polarity value of <0, then it is classified as a text with a negative sentiment.

3.4.1. Modelling

The BERT implementation stage consists of pre-processing, labeling, dataset splitting, sentiment classification, pre-training, fine-tuning, training, and evaluation. BERT implementation consists of data preparation and class initialization. The data Preparation dataset must be adjusted to the input received by BERT before training on BERT. BertTokenizer is needed to tokenize sentences

and generate appropriate input. This is done because BERT uses a specific vocabulary, depending on the model used. A tokenizer compiles sentences representing the input in BERT at the load model stage. Next, perform class initialization on the sentiment dataset document and sentiment data loader document on the class initialization stage. When the data training process is carried out, the data set will be divided into three categories: training, validation, and testing. The aim of dividing the dataset into three categories is so that machine learning knows the purpose of each dataset.
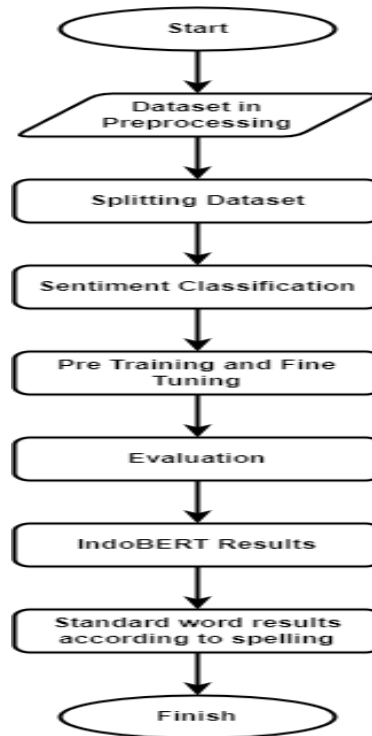

Figure 6. Implementation IndoBERT

BERT will go through a fine-tuning process by adjusting hyperparameters when the training stage begins. The fine-tuning process is done to find a configuration that works on the dataset. Below are the hyperparameters used in the training process:
1. The batch size is the number of samples fed into the network before the weights are adjusted. The larger the batch size, the longer it takes to complete one batch.
2. The epoch is the number of times the network looks at the entire data set. An epoch occurs when all instances have passed through the network, either forward pass or backward pass.
3. Learning level is one of the training parameters for calculating the weight correction value during the training process. The greater the learning level value, the faster the training process will run. Following are the hyperparameters recommended for fine-tuning that indicate optimal performance for specific tasks, namely:
    1. Batch Size: 32
    2. Level of learning: 2e-5
    3. Epoch: 10

The stage for determining the optimizer used is as follows,

```
import torch.optim as optim

# Tentukan optimizer
optimizer = optim.Adam(model.parameters(), lr=2e-5)
model = model.cuda()
```

Figure 7. The Stage for determining the optimizer

Figure 7 shows that the optimizer is used to optimize functions on the dataset. The optimizer used is Adam, with a learning rate of 2e-5. The number of epochs used is 10. The epoch is the number of times the network can view the entire dataset. Training is carried out with two datasets: the training data set and the validation dataset, as shown in Figure 7.
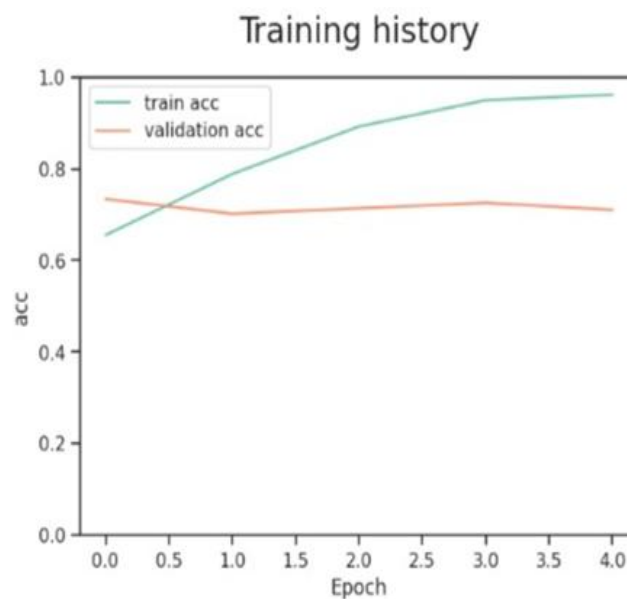

Figure 8. Training History

In Figure 8 above, the accuracy of the training data model is better than that of the validation data model. The learning curve shows increased accuracy results obtained during training data. In the validation model, the accuracy results are high but tend to decrease during the process. After training through the fine-tuning stage, the dataset is tested on sample sentences to prove that the predictions are correct.

3.4.2. Evaluation

This evaluation stage will evaluate the model created to determine to what extent the model can be classified. This evaluation uses a confusion matrix evaluation. This method is used in classification models to visualize and calculate the number of correct and incorrect predictions made by the model. Figure 6 below is the result of the resulting confusion matrix. The confusion matrix diagram is used to measure machine learning classification. At this stage, we discuss the results of two comparisons between validation data and testing data. The following are two comparisons between validation data and testing data.

1.   Validation Data
     This stage will discuss the validation results to show the confusion matrix by measuring how many models succeeded in predicting sentiment from the validation data. The results of the validation data confusion matrix can be seen in Figure 8. From the results of the validation confusion matrix in Figure 9, it can be seen that as many as 297 are in the True Positive (TP)

category, namely comments that are positive with the predicted results showing true positive, 22 data are included in the True Neutral (TNt) category, namely comments that have neutral value with the predicted results showing that they are truly neutral. One hundred data points are included in the True Negative (TN) category: comments with a negative value and the predicted results showing a true negative.
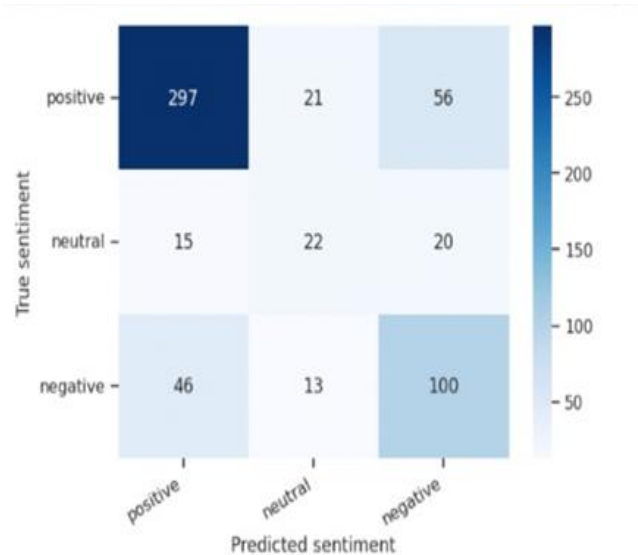

Figure 9. Confusion Matrix Diagram for Validation Data

The results of the F-1 score accuracy value from the validation data reached 71%, the macro AVG value from precision reached 60%, and the macro AVG value from the recall value reached 60%, as shown in Table 5.

Table 5. Results or Validation Data Accuracy Report

Result	Precision	Recall	F1-Score	Support
Positive	0.83	0.79	0.81	374
Neutral	0.39	0.39	0.39	57
Negative	0.57	0.63	0.60	159
Accuracy			0.71	590
Macro Avg	0.60	0.60	0.60	590
Weighted Avg	0.72	0.71	0.71	590

The macro average value of the F-1 score reaches 60%. After completing trials on validation data, the next step is to carry out trials testing data.

2. Data Testing
The final stage will cover the testing results to demonstrate the confusion matrix by measuring the number of models that successfully predict sentiment based on testing data. Figure 10 shows that 144 data are included in the True Positive (TP) category, namely positive comments, with the predicted results showing true positive. 13 data are included in the True Neutral (TNt) category, namely comments with a neutral value, with prediction results showing true neutral. Sixty data are included in the True Negative (TN) category, namely negative remarks, with the predicted results showing true negatives.
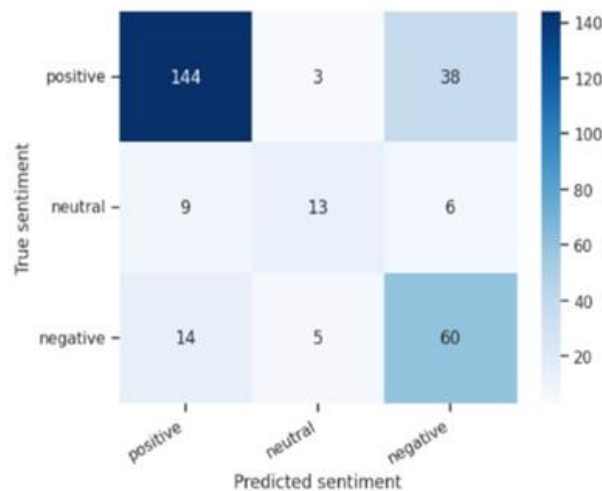
Figure 10. Confusion Matrix Diagram for Testing Data

Thus, it can be seen that the system can classify positive, neutral, and negative comments correctly. The confusion matrix is used to measure the performance of machine learning classification. The results of the confusion matrix for the testing data can be seen in Table 6.

Table 6. Confusion Matrix

Result	Precision	Recall	F1-Score	Support
Positive	0.86	0.78	0.82	185
Neutral	0.62	0.46	0.53	28
Negative	0.58	0.76	0.66	79
Accuracy			0.74	292
Macro Avg	0.69	0.67	0.67	292
Weighted Avg	0.76	0.74	0.75	292

Precision aims to calculate the percentage of inputs detected by the system. The precision value in the validation results is 83%. Recall seeks to calculate the percentage of inputs correctly identified by the system. The Recall value obtained in the validation is 79%. The accuracy of the validation test is 71%, which indicates the percentage of inputs correctly predicted by the neural network. F1-Score is the average obtained from precision and recall. The F1-Score in the validation testing results is 81%, which can be seen in Figure 9 and Figure 10. The precision value in the testing results is 86%. Recall in testing is 78%, and Accuracy in the testing phase has a value of 74%. This value of 74% is due to the limited and less complex training data used. If the training data used becomes more complicated, it will result in an increased accuracy value. The F1-Score is 82%, as can be seen in Table 7. Compared to previous research using the BERT algorithm, the F1-Score value in this research is higher, precisely 82%. The F1 score is 82%, which means that the majority of Indonesian people do not support LGBT. The neutral classification results perform significantly less than other labels in validation and testing. This is because the training data needs to sufficiently represent the variations within the neutral category, resulting in the model being poorly trained to recognize patterns related to this category.

Table 7. Result

Result	Precision	Recall	Accuracy	F1-Score
Validation Test	83 %	79%	71%	81%
Testing	86 %	78%	74%	82%

4. Conclusion

This research used the IndoBERT method in sentiment analysis of YouTube comments about LGBT and was successfully carried out well. The results show that the majority of Indonesian people do not support LGBT. IndoBERT's accuracy level in classifying con and pro sentences is 74%, with a precision value for positive words of 86%, a recall of 78%, and an F1 score of 82%.

References

[1] Y. E. Hoiriyah, N. Qomariyah, A. K. Darmawan, M. Walid, "Sentiment Analysis on LGBT issues in Indonesia with Lexicon-Based and Support Vector Machine Algorithms," *Pilar Nusa Mandiri : Journal of Computing and Information System*, vol. 19, n, no. No. 1, p. P. 27-36, 2023.

[2] Tomi Satalar; Titik Hidayati, "Permasalahan yang Berhubungan dengan Lesbian, Gay, Bisexual, dan Transgender: Literatur Review," *Jurnal Keperawatan Jiwa (JKJ) Persatuan Perawat Nasional Indonesia*, vol. Vol 9, No, 2021, doi: 10.26714/jkj.9.1.2021.213-220.

[3] F. F. Abdulloh and I. R. Pambudi, "Analisis Sentimen Pengguna Youtube Terhadap Program Vaksin Covid-19," *CSRID (Computer Science Research and Its Development Journal)*, vol. 13, no. 3, p. 141, Nov. 2021, doi: 10.22303/csrid.13.3.2021.141-148.

[4] L. Parabhoi and P. Saha, "Sentiment Analysis of YouTube Comments on Koha Open Source Software Videos," *International Journal of Library and Information Studies,* vol. 8, No. 4, pp. 96–102, 2018.

[5] N. P. V. D. Saraswati, N. Yudistira, and P. P. Adikara, "Analisis Sentimen terhadap Perundungan Siber pada Twitter menggunakan Algoritma Bidirectional Encoder Representations from Transformer (BERT)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 2, pp. 909–916, Mar. 2023, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12345.

[6] C. A. Putri, "Analisis Sentimen Review Film Berbahasa Inggris Dengan Pendekatan Bidirectional Encoder Representations from Transformers," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 6, no. 2, pp. 181–193, Jan. 2020, doi: 10.35957/jatisi.v6i2.206.

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." in *Proceedings of NAACL-HLT*, 2019, pp. 4171–4186.

[8] A. Elhan, M. K. D. Hardhienata, Y. Herdiyeni, S. H. Wijaya, and J. Adisantoso, "Analisis Sentimen Pengguna Twitter terhadap Vaksinasi COVID-19 di Indonesia menggunakan Algoritme Random Forest dan BERT," *Jurnal Ilmu Komputer dan Agri-informatika (JIKA)*, vol. 9, no. 2, pp. 199–211, Nov. 2022, doi: 10.29244/jika.9.2.199-211.

[9] A. Anjani, "ANALISIS SENTIMEN KAUM LGBT PADA MEDIA SOSIAL TWITTER MENGGUNAKAN ALGORITMA NAÏVE BAYES," *JTINFO : Jurnal Teknik Informatika*, vol. 1, no. 2, pp. 1–8, 2022, doi: 10.02220/jtinfo.v1i2.259.

[10] N. T. L. Mustakim, M. Ridwan, "Implementasi Algoritma Support Vector Machine untuk Analisis Sentimen LGBT di Indonesia," *Bulletin of Artificial Intelligence*, vol. 1, no. 2, pp. 36–42, 2022, [Online]. Available: https://journal.grahamitra.id/index.php/buai/article/view/31.

[11] T. Muhayat, A. Fauzi, and J. Indra, "Analisis Sentimen Terhadap Komentar Video Youtube Menggunakan Support Vector Machines," *Progresif : Jurnal Ilmiah Komputer*, vol. 19, no. 1, pp. 231–240, 2023, [Online]. Available: http://ojs.stmik-banjarbaru.ac.id/index.php/progresif/article/view/1060.

[12] R. Asrianto and M. Herwinanda, "Analisis sentimen kenaikan harga kebutuhan pokok dimedia sosial youtube menggunakan algoritma support vector machine," *Jurnal CoSciTech (Computer Science and Information Technology*, vol. 3, no. 3, pp. 431–440, Dec. 2022, doi: 10.37859/coscitech.v3i3.4368.

[13] A. R. I. B. Kurniawan, A. A. Aldino, "Sentimen Analisis Terhadap Kebijakan Penyelenggara Sistem Elektronik (PSE) Menggunakan Algoritma Bidirectional Encoder Representation From Transformers (BERT)," *Jurnal Teknologi dan Sistem Informasi*, vol. 3, no. 4, pp. 98–106, 2022, doi: 10.33365/jtsi.v3i4.2204.

[14] S. Dharmawan, V. C. Mawardi, and N. J. Perdana, "Klasifikasi Ujaran Kebencian Menggunakan Metode FeedForward Neural Network (IndoBERT)," *Jurnal Ilmu Komputer*

dan Sistem Informasi (JIKSI), vol. 11, no. 1, Jun. 2023, doi: 10.24912/jiksi.v11i1.24066.

[15] EDWINA AYU CHRISTY, "Youtube comments on LGBT in Indonesia," 2023. https://www.kaggle.com/datasets/edwinaayuchristy/youtube-comments-on-lgbt-in-indonesia.

[16] T. N. Ph.D and A. Amalanathan, "DATA PREPROCESSING IN SENTIMENT ANALYSIS USING TWITTER DATA," *International Educational Applied Research Journal (IEARJ),* vol. 3, pp. 89–92, 2019.

[17] H. Jayadianti, W. Kaswidjanti, A. T. Utomo, S. Saifullah, F. A. Dwiyanto, and R. Drezewski, "Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN," *ILKOM Jurnal Ilmiah*, vol. 14, no. 3, pp. 348–354, Dec. 2022, doi: 10.33096/ilkom.v14i3.1505.348-354.

[18] DOsinga and D. Osinga, "Deep Learning Cookbook Notebooks." https://github.com/DOsinga/deep_learning_cookbook.

[19] A. Tharwat, "Classification assessment methods," *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, Jan. 2021, doi: 10.1016/j.aci.2018.08.003.

[20] J. Brzozowska, J. Pizoń, G. Baytikenova, A. Gola, A. M. Zakimova, and K. Piotrowska, "DATA ENGINEERING IN CRISP-DM PROCESS PRODUCTION DATA – CASE STUDY," *Applied Computer Science*, 2023, [Online]. Available: https://api.semanticscholar.org/CorpusID:264533715.