Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

Sentiment Analysis Of Indihome Service Based On Geo Location Using The Bert Model On Platform X

Robiatul Adawiyah Siregar [1], Fitriyani [*2], Lazuardy Syahrul Darfiansa[3]

[1,2,3]Informatics, Engineering Faculty, Telkom University, Indonesia

Email: [2]fitriani@telkomuniversity.ac.id

Received : Jun 30, 2025; Revised : Aug 23, 2025; Accepted : Sep 2, 2025; Published : Oct 16, 2025

Abstract

The rapid growth of internet usage in Indonesia has led more people to express their feelings, whether positive or negative, about online services, including IndiHome, through social media platforms such as X (formerly Twitter). This study aims to analyze public sentiment toward IndiHome services based on geographic location using the IndoBERT natural language processing model. The data consists of 3.307 Indonesian tweets that are geo-tagged and categorized into three sentiment groups: good, okay, and bad. The research process involves collecting the data, cleaning it (organizing and splitting words), and testing the IndoBERT model with a confusion matrix and classification scores. The findings reveal that negative feelings are more prevalent in most locations, especially in Java. The IndoBERT model achieved its highest accuracy of 80% in detecting negative sentiment. However, there is still room for improvement in distinguishing between positive and neutral sentiments, possibly due to data imbalance. The study shows how combining sentiment analysis with geo-location can provide strategic insights to service providers. In practical terms, these insights can help IndiHome prioritize infrastructure upgrades, improve customer support in areas with high dissatisfaction, and assist policymakers in promoting fairer digital access across regions. Beyond these practical implications, this study also contributes to the field of informatics by demonstrating the application of a transformer-based NLP model (IndoBERT) combined with geo-tagged data for regional sentiment mapping- a relatively unexplored approach in the Indonesian context. The integration of geospatial analysis with sentiment classification offers methodological advances for NLP-based service evaluation beyond business applications.

Keywords : *Geo Location, IndiHome, IndoBERT, Platform X, Sentiment Analysis.*

1. INTRODUCTION

The increasing internet penetration in Indonesia has changed the way people use digital services, including fixed broadband services. IndiHome, a subsidiary of PT Telekomunikasi Indonesia, offers triple-play services, including internet, telephone, and IPTV, making it one of the leading Internet Service Providers (ISPs) in Indonesia [1]. With a market share of around 8.7%, this service remains the primary choice for Indonesian users [2]. However, public feedback on social media platforms, especially X (formerly Twitter), reflects a variety of opinions about the service, both positive and negative. This feedback is often geo-tagged, providing a rich source of data that can be used to understand regional satisfaction and improve service quality [3].

Sentiment analysis, also known as opinion mining, is an important technique for extracting subjective information from text. It categorizes opinions into positive, negative, and neutral, and is widely used in areas such as customer service, product reviews, and political analysis. In the field of Internet Service Providers (ISPs), sentiment analysis can help service providers gain insight into customer perceptions and identify service deficiencies [4]. Recent advances in natural language translation (NLP), particularly deep learning models such as BiDirectional Encoder Representations from Transformers (BERT), have significantly improved the accuracy and efficiency of sentiment

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

classification. BERT is a pre-drilled Transformer model that understands context from both directions, making it ideal for analyzing sentiment in short, informal texts such as tweets. IndoBERT has been optimized for Indonesian, resulting in high accuracy for translating local languages [5].

Although many research efforts have used traditional machine learning methods like Naive Bayes and Random Forest for sentiment analysis, their performance tends to be limited when dealing with complex linguistic structures [6]. In contrast, models based on BERT demonstrate better understanding of context and higher accuracy in classification [7]. However, there is a lack of research focused on sentiment analysis using BERT within the field of geo-tagged data related to Indonesian Internet service providers. Recent studies have shown the effectiveness of IndoBERT and hybrid deep-learning architectures in capturing sentiment patterns from short, noisy Indonesian texts [8]. Additionally, the use of geospatial analysis in sentiment research remains limited, although early efforts in disaster mapping [9] and election monitoring [10] are promising.

However, research on the use of BERT for location based sentiment analysis in Indonesia remains limited. Therefore, this study aims to assess the effectiveness of the IndoBERT model in classifying customer sentiment toward IndiHome services by incorporating geolocation data. Specifically, the objectives include developing an IndoBERT based sentiment classification model for geo-tagged tweets related to IndiHome, analyzing the distribution of positive, negative, and neutral sentiments across different regions in Indonesia, and identifying areas with low customer satisfaction to inform local service improvements. This approach is expected to enable the model to effectively capture public opinions from various regions and support strategic decision making in service improvement.

The geographical location of users significantly affects how they perceive the quality of IndiHome services. Users in Java, where digital engagement and service expectations are high, tend to report more dissatisfaction compared to those in Sumatra and Sulawesi, where infrastructure development and population density are lower. Previous studies have demonstrated that regional differences in infrastructure directly influence user satisfaction with internet services. Despite numerous studies applying sentiment analysis to Indonesian texts, only a few have used geo-tagged data for service evaluation, making this approach relatively new researchch.

2. RESEARCH METHOD

The implementation of this study, illustrated in Figure 1, consists of a series of stages designed to perform sentiment analysis of IndiHome services by incorporating geo-location data using the BERT model on Platform X. The process begins with collecting user generated content from Twitter using relevant keywords. The collected data then undergoes a preparation phase, including text cleaning, normalization, tokenization, and stemming, to ensure it is ready for analysis. Finally, the tweets are sentiment labeled (positive, neutral, negative) and grouped based on the users' geographic locations for further interpretation.

2.1.1. Problem Identification

Although social media platforms like Platform X (Twitter) have become the primary avenue for people to share their opinions on digital services such as IndiHome, most previous research remains limited to general sentiment analysis without taking into account the geographical location of users. In reality, the quality of internet services is heavily affected by regional conditions, infrastructure, and user density in each area. Additionally, traditional algorithms like Naïve Bayes or Random Forest are seen as less effective at understanding the complex linguistic contexts in Indonesian. Hence, a new approach is necessary that can classify public sentiment more accurately while linking it to location data. This study addresses this gap by employing the IndoBERT model and geo-tagged data to assess public perceptions of IndiHome services across different regions Indonesia.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

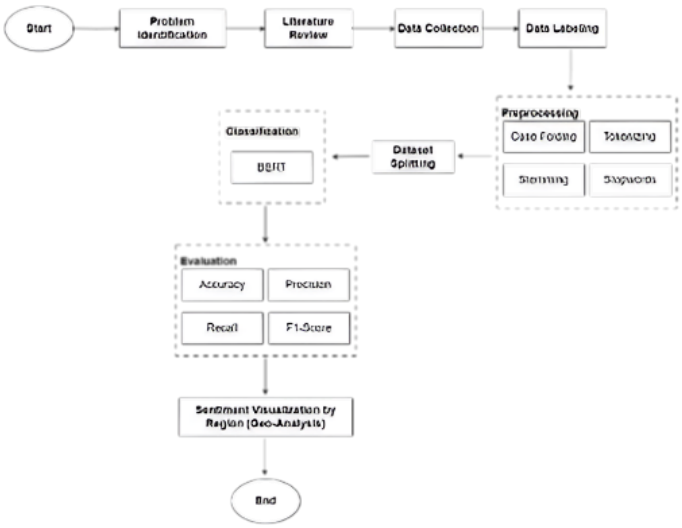

Figure 1. Research methodology.

2.1.2. Problem Identification

Although social media platforms like Platform X (Twitter) have become the primary avenue for people to share their opinions on digital services such as IndiHome, most previous research remains limited to general sentiment analysis without taking into account the geographical location of users. In reality, the quality of internet services is heavily affected by regional conditions, infrastructure, and user density in each area. Additionally, traditional algorithms like Naïve Bayes or Random Forest are seen as less effective at understanding the complex linguistic contexts in Indonesian. Hence, a new approach is necessary that can classify public sentiment more accurately while linking it to location data. This study addresses this gap by employing the IndoBERT model and geo-tagged data to assess public perceptions of IndiHome services across different regions Indonesia.

2.1.3. Literature Review

Research on sentiment analysis has advanced considerably alongside the growing use of social media as a platform for expressing public opinion. In this context, various approaches have been adopted, ranging from traditional classification techniques to models employing deep learning. One popular model with excellent performance is BERT (Bidirectional Encoder Representations from Transformers), developed by Google, renowned for its ability to understand word context bidirectionally within a single sentence.

The study by Fatma Sjoraida.dkk [11] utilised BERT to evaluate public opinions on the movie Dirty Vote. This model achieved impressive results, with accuracy of 85%, precision of 86%, recall of 84%, and an F1-score of 85%. These outcomes demonstrate BERT's effectiveness in classifying Indonesian texts, particularly in the realm of public opinion often expressed on social media.

Similar research conducted by A. Elhan, dkk [12] compared the performance of Random Forest and BERT algorithms in analysing public attitudes towards COVID-19 vaccination in Indonesia using Twitter data. BERT exhibited superior performance, with an accuracy of 82% and an F1-score of 79%, whereas Random Forest recorded 81% accuracy and an F1-score of 74%. These findings suggest that transformer based methods, including BERT, are better suited to capturing the depth and nuance of informal texts such as tweets, which frequently contain abbreviations, symbols, and nonstandard sentence structures.

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

Meanwhile, [13] applied the BERT model to analyse sentiments in user reviews of the Ruang Guru educational app on the Google Play Store. Their model achieved 99% accuracy and an F1-score of 98.9%, further affirming BERT's reliability in processing short Indonesian texts. This effectiveness arises from BERT's pre-trained capacity to grasp word context in a bidirectional manner, enabling more precise understanding of word meanings based on surrounding text.

In addition, prior research by Putri and her team utilised the Naïve Bayes algorithm to classify sentiments related to public services and political issues. Although simple and efficient, this approach has limitations in understanding context and complex word meanings, with accuracy typically ranging from 75% to 80%.

However, most previous studies employing BERT and other classification methods face significant limitations, especially in sentiment mapping concerning users' geographic locations (geo-sentiment analysis). Usually, these studies focus solely on sentiment classification based on overall content, without considering spatial factors like the user's location, which can provide a deeper understanding of service quality.

In the realm of digital services such as IndiHome, users' experiences and viewpoints are heavily influenced by geographical factors. The user's geographic location plays a crucial role in determining the quality of their consumer experience, including elements like network density, internet infrastructure condition, and the number of active users in an area. Those residing in regions with robust digital infrastructure, such as Java Island—the hub of urbanisation and digital activity in Indonesia—tend to have higher expectations for service quality. This often results in increased complaints submitted via social media when encountering issues.

Building on prior research that has mainly focused on sentiment classification without considering users' geographic locations, this study introduces a new approach by combining IndoBERT based sentiment analysis with geo-tagged data to evaluate perceptions of IndiHome services. This method not only confirms BERT's effectiveness in classifying informal Indonesian text but also advances its use in spatially monitoring service quality, aiding regional service providers in decision making.

2.1.4. Data Collection

Data collection in this research involves two main methods, namely literature study and data crawling [14]. The literature study aims to gather comprehensive information from journals, articles, and digital sources to build a strong theoretical foundation. For data collection, tweets were retrieved from Platform X (formerly known as Twitter) by using tools such as Tweet Harvest or scraping methods with Twitter authentication tokens. A total of 3,307 tweets related to IndiHome services were retrieved on October 25, 2024, using keywords such as "IndiHome", 'slow', 'down', and "smooth". The tweets were saved in CSV format for use in further analysis. To provide a deeper understanding of the data structure applied in this study, a detailed breakdown of each element in the dataset is presented in Table 1.

Table 1. Feature description of the selected dataset.

Attributes	Details
Username	Feature description of the selected dataset.
Screen Name	A unique ID for the user's display name.
TweetAt	The date the tweet was posted.
Original Tweet	The original text content of the tweet.
Sentiment	Sentiment class labels (three categories).
Locations	The location of the user when creating the tweet

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

2.1.5. Data Labeling

This research uses a dataset of tweets categorized into three emotion groups negative (-1), neutral (0), and positive (1). A total of 3,307 tweets were collected from Platform X using a crawling process. After gathering the data, it was divided into three parts training, validation, and testing. The splitting was done in two steps with stratified sampling to keep the proportions of each emotion category consistent. Consequently, 70% of the data was allocated for training, 20% for validation, and 10% for testing. The distribution of these categories is shown in Figure 2.

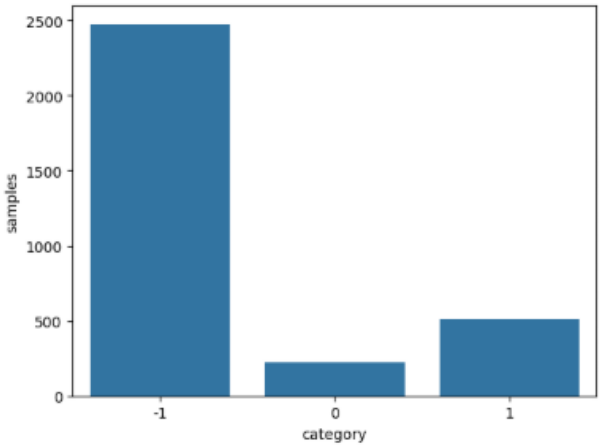

Figure 2. Class distribution

2.1.6. Preprocessing

In sentiment classification tasks based on short texts like tweets, the order and selection of preprocessing steps significantly influence model performance. This study applies a common set of preprocessing steps, including case folding, cleaning, tokenisation, normalisation, stopword filtering, and stemming. This set of preprocessing is developed based on widely used practices in natural language processing (NLP) and tailored to the characteristics of data from Indonesian language social media. All steps are carried out using Google Colab with Python scripts, starting with converting text to lowercase and removing foreign characters, numbers, punctuation marks, emojis, and links. Table 2 shows the preprocessing resultsesults.

Table 2. *Preprocessing*

Before	After
Signal indihome gembel bgt ya. Udh byr mahal kualitas pret @IndiHomeCare	signal indihome gembel banget ya udah bayar mahal kualitas pret @indihomecare

Tokenization and normalization steps are used to simplify nonstandard word forms or abbreviations that frequently appear on social media. Subsequently, filtering out empty words and restoring words to their original form are performed. This preprocessing was done to create clean and consistent data before it was used as input for the BERT model. Since the source data came from Platform X, which is usually disorganized and contains many informal language variations, this step was necessary to improve data quality and reduce distractions that could hinder the classification

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

process. Assessment of the model demonstrated notable improvements in F1 scores and accuracy, especially in distinguishing between neutral and negative tweets. These preprocessing steps align with established practices in Indonesian NLP, as outlined in the IndoNLU benchmark [15], which emphasizes how normalization and tokenization significantly affect IndoBERT performance.

Once the text preprocessing is finished, the dataset must be formatted to suit the BERT model's requirements. To do this, a tokenisation step is carried out using a tokenizer specific to the model. This tokenisation adds a [CLS] token at the start of the sentence to denote it as input for classification, followed by a [SEP] token at the end as a separator, and a [PAD] token if the sentence is shorter than the predefined maximum length [16]. If the sentence exceeds this length, it will be truncated. In this study, the maximum length is set at 512 tokens based on the analysis of text length distribution within the dataset. The tokens are then converted into numerical input IDS and complemented with an attention mask to highlight the input parts that should be attended to by the model . Tokens outside the vocabulary are broken into smaller segments using the WordPiece technique . This process is executed using the encode_plus() function of the indobenchmark/indobert-base-p1 tokenizer, enabling the input to be used during BERT's fine-tuning stage for sentiment classification tasks.

2.1.7. Dataset Splitting

The process of splitting the dataset is done in two main steps to ensure fair and separate model assessment. In the first step, the data is split into 70% for the training process and 30% as temporary data. Subsequently, the 30% temporary data is further divided, with 67% allocated to validation data and 33% to testing data. Thus, the final proportion of data is 70% for training, 20% for validation, and 10% for testing. This method utilizes stratification techniques to keep the class distribution consistent across each data subset.

In addition to using the standard 70:20:10 split, this study also explored several alternative ratios such as 80:20, 90:10, and 30:70 to assess the performance of the model under varying proportions of training and testing data. These variations in ratios serve to test the extent to which the model can adapt and generalize to different dataset sizes.

2.1.8. Fine-Tuning

Fine-tuning is the process of adjusting a pre-trained model to better fit a specific task, such as emotion preference classification. In this study, the indobenchmark/indobert-base-p1 model derived from Hugging Face is fine-tuned to perform classification with three sentiment categories (positive, neutral, negative) regarding IndiHome services using data from platform X [15].

Before undergoing fine-tuning, the data first goes through a tokenization process, the addition of special tokens [CLS] and [SEP], and conversion to input_ids and attention_mask [15]. Once the data has been processed, it is fed into the BERT model and proceeds to the classification stage using a fully connected layer and sigmoid activation function.

For the fine-tuning process to provide the best performing model, it is important to have accurate hyperparameter settings. The three main elements that most affect the performance of the model are batch size, learning rate, and number of epochs. Determining the best values is done through a series of iterative tests such as grid search or trial and error. To optimally support the fine-tuning process, adjustments were made to several key hyperparameters that strongly influence the performance of the model. These values were determined based on exploratory results and iterative testing. Full details of the hyperparameter configuration used in this study can be seen in table 3.

Hyperparameters were chosen based on evidence from previous IndoBERT fine-tuning experiments. A learning rate of 2e-5 is standard for transformer-based fine-tuning to ensure training stability [17], while a batch size of 16 was selected due to GPU memory limitations. The epoch limit of

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

10 was set to prevent underfitting, but since the validation loss indicated overfitting from the 4th epoch onwards, early stopping should be considered in future work.

Table 3. *Hyperparameters*

Hyperparameter	Value
Max sequence length	512
Batch size	16
Optimizer	AdamW
Learning rate	2e-5
Epoch	10

Recent studies also suggest that incorporating dropout regularization during IndoBERT fine-tuning can reduce overfitting and improve model generalization [18]. Additionally, using geo-tagged tweets may introduce bias because not all users enable location sharing. This could lead to overrepresentation of urban areas such as Java, restricting the applicability of the findings to rural areas regions.

2.1.9. Model Classification

This study uses the IndoBERT model to classify public sentiment toward IndiHome services based on tweet data collected from Platform X. IndoBERT, a transformer based language [19] model pre-trained specifically for the Indonesian language, is fine-tuned to perform multi class sentiment classification with three categories positive, neutral, and negative. Before classification, the text data undergoes a preprocessing pipeline that includes tokenization and formatting to meet BERT's input requirements. After fine-tuning, the model's performance is evaluated using a confusion matrix, which shows the number of correct and incorrect predictions across classes. The matrix features four key indicators true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). These metrics are used to calculate overall performance measures such as accuracy, precision, recall, and F1-score [17]. The results indicate that IndoBERT performs well in detecting negative sentiment, demonstrating high precision and recall scores, but is less effective at identifying neutral and positive sentiments. This limitation seems to stem from class imbalance and contextual ambiguity within the dataset. Overall, the evaluation confirms IndoBERT's ability to classify sentiment, while also emphasizing the need for further refinement, especially in handling underrepresented sentiment classes categories.

2.1.10. Visualization of Sentiment by Geographic Location

The geolocation visualization stage aims to display the results of sentiment analysis based on user location in a clear and visual way. This process shows the distribution of positive, neutral, and negative sentiment types across different specific areas, allowing the geographic pattern of public opinions on IndiHome services to be observed [20]. This visualization helps researchers and related parties understand sentiment trends in each location, supporting more accurate decision-making to improve service quality. Sentiment analysis based on geolocation reveals a strong dominance of negative sentiment, especially in Java. Out of 1,767 tweets from Java, 1,352 were classified as negative—far more than neutral or positive sentiments. The same pattern was observed in Sulawesi and Sumatra, although with fewer tweets.

3. RESULT

The BERT (Bidirectional Encoder Representations from Transformers) model is trained using data from platform X (formerly Twitter). The dataset includes various tweets classified into three sentiment

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

categories negative, neutral, and positive. The model employed is IndoBERT, a pre-trained version of BERT tailored for the Indonesian language, utilising the Indo4B dataset gathered from sources such as social media, blogs, and websites.

In the preprocessing stage, tweets are tokenised using the encode_plus() function of the BERT tokenizer, generating input_ids (numeric representations of the text) and attention_mask (indicators for tokens to consider). The data is further cleaned through normalization and filtering to remove irrelevant characters, symbols, and stopwords. During fine-tuning, the pre-trained IndoBERT model is adapted by adding a classification layer for sentiment analysis. The chosen hyperparameters include a batch size of 16, a learning rate of 2e-5, and the training process is conducted for 10 epochs.

3.1.1. Data Collection

A total of 3,307 tweets about IndiHome services were gathered from Platform X (formerly Twitter) using keywords like “IndiHome”, “slow”, “down”, and “smooth”. These tweets were geo-tagged and spread across Indonesia’s main islands, with most coming from Java.

Tabel 4. Dataset Distribution by Region

Dataset Distribution by Region				
Region	**Negative (-1)**	**Neutral (0)**	**Positive (1)**	**Total**
Java	1,352	129	286	1,767
Sumatra	143	12	20	175
Sulawesi	100	8	26	134
Total	1,595	149	332	2,076

This distribution shows a high prevalence of negative tweets, especially in Java, which indicates greater dissatisfaction in areas with dense digital activity infrastructure.

3.1.2. Preprocessing

The preprocessing stage involved case folding, stopword removal, normalization, tokenization, and stemming. This process helped eliminate noise caused by informal language typical in tweets.

Table 5. *Preprocessing*

Before	**After**
Signal indihome gembel bgt ya. Udh byr mahal kualitas pret @IndiHomeCare	signal indihome gembel banget ya udah bayar mahal kualitas pret indihomecare

This cleaning ensures that tweets are formatted uniformly for the IndoBERT tokenization process. Next, tokenization organizes words into subword units using the WordPiece method to match IndoBERT vocabulary.

3.1.3. Evaluation Model

The performance evaluation of the BERT model in sentiment clustering was also conducted by examining the confusion matrix. In general, the model obtained an accuracy rate of up to 80%, which shows a satisfactory performance. However, to better understand the model's ability to identify each sentiment category, confusion matrix analysis is crucial.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

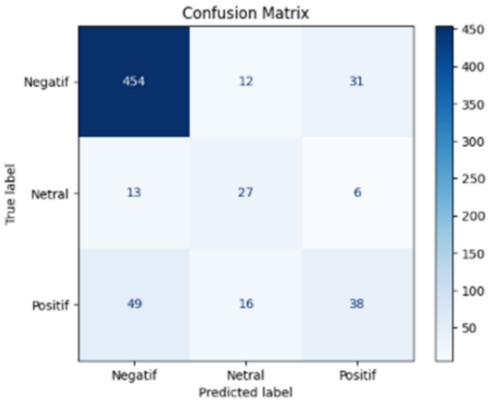

Figure 3. Class distribution

In Figure 3, it can be seen that the BERT model performed best in classifying negative sentiments with 454 accurate predictions, although there were some misclassifications into neutral (12) and positive (31) categories. For neutral sentiment, the model performed below expectations with only 27 correct predictions, while 13 instances were categorized as negative and 6 as positive. This indicates that the model faced challenges in distinguishing neutral nuances, possibly due to the overlapping meanings between the classes. Within the positive class, only 38 instances could be accurately predicted, while 49 were erroneously classified as negative and 16 as neutral. This underperformance in the neutral and positive categories suggests that the model has difficulty in understanding subtle differences in context or perhaps due to an imbalance in the distribution of data between the classes.

Tabel 6. Classification Report

Classification Report				
Category	**Precision**	**Recall**	**F1-Score**	**Support**
Negatif	0.88	0.91	0.90	497
Netral	0.49	0.59	0.53	46
Positif	0.51	0.37	0.43	103

Based on the evaluation results presented in Table III, the IndoBERT model shows excellent performance in identifying tweets containing negative sentiments, with a precision value of 0.88, recall 0.91, and F1-score 0.90 from a total of 497 samples. This shows that the model can consistently and correctly recognize negative sentence patterns. In contrast, the model performance for the neutral and positive categories was low. For the neutral category, the precision and recall values are 0.49 and 0.59 respectively, resulting in an F1-score of 0.53 from a total of 46 samples. Meanwhile, for the positive category, the model only achieved a precision of 0.51 and recall of 0.37, with an F1-score of 0.43 out of 103 available samples. The low performance in these two categories may be due to the imbalance in the data distribution, as well as context confusion in the expression of neutral and positive tweets.

Overall, the model achieved 80% accuracy on 646 test data, with a macro average F1-score of 0.62 and a weighted average F1-score of 0.80. This imbalance in performance between classes suggests the importance of implementing strategies to balance the data or improve the feature representation, in order for the model to improve its generalization ability to all sentiment classes more proportionally.

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

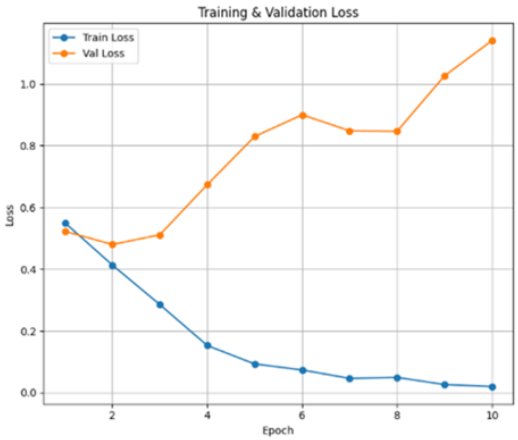

Figure 4. Visualization results of training loss and validation loss curves.

Model training was carried out for 10 epochs, with indicators that the training loss decreased continuously, indicating a successful learning process on the training data. However, the validation loss started to increase after the 4th epoch, indicating overfitting. This suggests that the best performance of the model may occur before the 5th epoch, so it is necessary to apply methods such as early stopping or regularization to improve the generalization ability of the model.

Figure 5. Training Accuracy and Validation Accuracy Results.

The curves shown in Figure 5 show the changes in training accuracy and validation accuracy throughout the model's training sessions over 10 epochs. The accuracy on the training data shows a significant increase to almost 100%, indicating that the model can understand the patterns in the training data very well. On the other hand, the accuracy for the validation data tends to remain unchanged and remains at around 80%, without showing any significant improvement despite the increase in the number of epochs. In fact, there are variations that indicate the instability of the model's performance when dealing with new data. The large difference between training accuracy and validation accuracy further confirms the signs of overfitting, where the model overfits itself to the training data and sacrifices its ability to generalize. To address overfitting observed during training (as seen in Figures 5 and 6), methods such as early stopping, dropout regularization, and weight decay are recommended for future experiments. These methods have been shown to improve generalization in IndoBERT models [21].

3.1.4. Distribution of Sentiment by Geographic Region

Based on the analysis conducted, it is found that:

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

1. Java was the region with the most tweets, with 1,767 tweets, of which 1,352 were classified as negative sentiment. This figure is much higher than the number of neutral (129 tweets) and positive (286 tweets) sentiments in the area.
2. Sumatra Island came in second with a total of 175 tweets, consisting of 143 negative tweets, 12 neutral tweets and 20 positive tweets.
3. Sulawesi Island recorded the lowest number of tweets, 134 tweets, with 100 negative tweets, 8 neutral tweets and 26 positive tweets.

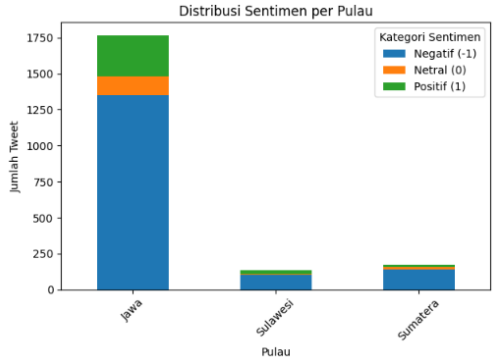

Figure 6. Inter island Distribution of IndiHome Service Sentiment.

Tabel 7. Sentiment Distribution Table By Island

Pulau	Negatif (-1)	Netral (0)	Positif (1)	Total
Jawa	1.352	129	286	1.767
Sumatera	143	12	20	175
Sulawesi	100	8	26	134

In Table IV, there is a description of the sentiment distribution of tweets related to IndiHome services segmented into three main regions in Indonesia, namely Java, Sulawesi and Sumatra. The analysis conducted reveals that negative sentiment dominates in all regions, with Java Island showing the highest number of 1,352 negative tweets out of a total of 1,767 tweets studied. This figure is far greater than the neutral (129) and positive (286) tweets coming from the same region. Sulawesi recorded 100 negative, 8 neutral and 26 positive tweets, while Sumatra reported 143 negative, 12 neutral and 20 positive. This pattern of dominance of negative sentiment suggests that people are more likely to be dissatisfied with the service under study, IndiHome.

The analysis of Sumatra and Sulawesi reveals that although the number of tweets is considerably lower than Java, negative sentiment still predominates. In Sumatra, 143 of 175 tweets were negative, indicating ongoing dissatisfaction even with smaller sample sizes. Likewise, in Sulawesi, 100 out of 134 tweets expressed negative sentiment, suggesting that issues with connectivity and infrastructure remain. These results show that dissatisfaction is not confined to the highly urbanised Java but extends to other islands, although to a lesser extent volume.

The high rate of negative sentiment in Java can be explained by various structural factors. As the center of digital activity and the region with the highest internet access rate in Indonesia, Java is an area where users have higher expectations of digital service quality. Increased urbanization, better education

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

levels, and awareness of consumer rights encourage Javanese to be more active in expressing their dissatisfaction online.

Figure 7. Negative Sentiment Word Cloud.

Figure 7 shows the word cloud visualization for negative sentiment obtained from the BERT model classification of tweets from IndiHome users. This representation features key words such as “internet”, “not”, “wifi”, ‘off’, and “interference”, reflecting the main complaints around inconsistent connections, service issues, and customers' dissatisfaction with technicians and billing. Geographically, these results are in line with the predominance of negative sentiments in regions such as Java, which has high service expectations. This word cloud provides significant thematic insights that service providers can leverage to focus more on improving quality and responding to customer complaints.

Figure 8. Positive Sentiment Word Cloud.

Figure 8 displays a word cloud that illustrates the positive sentiment from the BERT model classification of IndiHome users' tweets. The most frequently occurring words such as “internet”, “smooth”, “stable”, ‘wifi’, and “Telkomsel” indicate that user recognition is mostly related to good connections and responsive services. These findings suggest that network quality, ease of use, and good interactions are key elements that drive positive sentiment, so they can be focus areas for service providers to improve in different regions.

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

4. DISCUSSION

Overall, the IndoBERT model showed strong ability to classify user sentiments toward IndiHome services based on geo tagged tweets from Platform X (formerly Twitter). The model reached an overall accuracy of 80%, with especially high performance in detecting negative sentiment, indicated by a precision of 0.88, recall of 0.91, and F1-score of 0.90 [22].

Tabel 8. Comparison of Sentiment Analysis Performance in Previous Studies

Study	Dataset	Method	Accuracy	Notes
Tyas et al. (2022)	IndiHome Tweets	Naïve Bayes	~80%	Without geo-location
Elhan et al. (2021)	COVID-19 Tweets	Random Forest vs BERT	81% vs 82%	BERT superior
Fatma Sjoraida et al. (2024)	Dirty Vote Tweets	BERT	85%	Political domain
This Study (2025)	IndiHome Tweets (Geo-tagged)	IndoBERT	80%	integrate geo-location

As shown in table 8, the proposed IndoBERT model demonstrates comparable or even superior results compared to previous studies using conventional machine learning methods (Naive Bayes, Random Forest) and is nearly comparable to IndoBERT in other areas. The key innovation of this research lies in the integration of geographically relevant data, which provides previously unexplored region-based sentiment mapping.

These results confirm the model's effectiveness in identifying dissatisfaction patterns within short, informal texts like tweets, which are commonly used to express service complaints. Similar findings have been seen in previous studies involving transformer based models in sentiment analysis tasks, emphasizing BERT's contextual understanding of text in low resource languages such as Indonesian.[11].

However, the performance for neutral and positive sentiments is lower (F1-score of 0.53 and 0.43 respectively), which is likely due to the imbalance in the amount of data between classes and ambiguity in the meaning of tweets. This phenomenon is common in sentiment analysis on social media with uneven data distribution. Geolocation analysis reveals that Java Island recorded the highest dissatisfaction, with 1,352 out of 1,767 tweets classified as negative. Sumatra (143 negative tweets out of 175) and Sulawesi (100 negative tweets out of 134) also exhibited a predominance of negative sentiment, despite having smaller overall volumes. This suggests that dissatisfaction is not solely concentrated in Java, the digital hub, but is also present in other regions with comparatively limited infrastructure. These findings highlight that geographical factors, such as infrastructure readiness and population density, significantly influence public perception of IndiHome services. Service providers can use this insight to prioritise capacity upgrades in Java while also expanding infrastructure in Sumatra and Sulawesi to address service disparities across regions. The accuracy and loss curves show symptoms of overfitting, where training accuracy almost reaches 100% while validation accuracy stagnates [23]. This indicates that the model overfits the training data and is less able to handle new data, thus regularization or early stopping is required in future studies [24].

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

Future research can expand the dataset beyond Platform X by including customer service logs, survey results, or other social media platforms to enhance sentiment analysis. Real-time sentiment monitoring could also offer dynamic insights, allowing IndiHome to respond proactively to customer feedback dissatisfaction. Additionally, incorporating topic modeling techniques like Latent Dirichlet Allocation (LDA) would enable a more detailed analysis of specific service issues being discussed [25]. This indicates that the model overfits the training data and is less effective at handling new data, so regularization or early stopping will be necessary in future studies.

5. CONCLUSION

This study shows that the IndoBERT model effectively classifies user sentiment toward IndiHome services using data from Platform X (Twitter), with an accuracy of 80%. It performs well in detecting negative sentiment (precision 0.88, recall 0.91, F1-score 0.90), although results for neutral and positive sentiments are limited due to data imbalance and ambiguity in tweet context. Geo-location analysis revealed distinct regional differences: Java Island produced the highest number of negative tweets (1,352 out of 1,767), followed by Sumatra (143 out of 175) and Sulawesi (100 out of 134). This pattern indicates higher expectations and greater digital engagement in Java, while dissatisfaction in Sumatra and Sulawesi suggests that connectivity and infrastructure gaps remain issues beyond urban areas centers.

Compared to earlier studies such as those by [10], which used BERT for sentiment analysis on vaccine related tweets (accuracy 82%), or Tyas et al. [3], which applied Naïve Bayes to analyze IndiHome sentiment without location data (accuracy around 80%) this research advances previous work by combining IndoBERT with geo tagged data to more precisely identify regional dissatisfaction. This approach allows for more targeted recommendations for service providers. The findings emphasise that improvement strategies should not only target highly urbanised Java but also include infrastructure development in Sumatra and Sulawesi to ensure more equitable service quality across the regions Indonesia. This study makes a unique contribution by combining IndoBERT with geo-location analysis for sentiment classification of broadband services in Indonesia, an approach rarely explored in prior studies. Beyond practical recommendations for service providers, this study contributes to Informatics by introducing a framework that integrates transformer-based sentiment classification with geospatial analysis. This methodological contribution enables the mapping of digital service satisfaction at regional level, advancing NLP applications in under-researched contexts such as Indonesian broadband services. Such integration may also inspire future work in geo-tagged opinion mining for other domains like public health, education, and disaster response. Limitations include the imbalance of sentiment classes and the underrepresentation of rural areas due to geo-tagging bias. Future work should consider balancing techniques such as SMOTE, incorporating multiple data sources, and expanding to real-time analysis for service monitoring. Future research could build on this by addressing class imbalance with techniques like SMOTE, testing the model on more balanced datasets, and adding topic modeling (e.g., LDA) or real time sentiment tracking to enhance analysis and provide dynamic insights for service management improvement.

ACKNOWLEDGMENTS

The authors express their sincere appreciation to all those who have played a role in supporting this research. Thanks to Telkom University for providing facilities, direction, and a conducive academic environment in the implementation of this research. The authors also appreciate the contributions of the supervisors who have provided constructive input and valuable guidance throughout the process of preparing this research. My deepest gratitude also goes to my family and closest friends for their

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

continuous prayers, encouragement, and motivation. Hopefully, the results of this research can provide real benefits and contributions in the development of geo location based sentiment analysis in Indonesia.

The approach of combining IndoBERT and geolocation data is a new contribution that not only improves classification accuracy, but also helps understand service satisfaction spatially for the purposes of improving area based service strategies. However, this study has limitations, such as not considering external factors (e.g. infrastructure and regulations), inequality of sentiment data, and uneven regional representation. This limits the generalizability of the results to the whole of Indonesia. In the future, research can be extended with data balancing techniques such as SMOTE, topic integration using LDA, and real time sentiment monitoring. A multivariate approach is also recommended to improve the accuracy and practical usefulness of the model in service management.

REFERENCES

[1] R. Puspita and A. Widodo, "Analisis Sentimen terhadap Layanan Indihome di Twitter dengan Metode Machine Learning," vol. 6, no. 4, pp. 759–766, 2021, doi: 10.32493/informatika.v6i4.13247.

[2] R. Tineges, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen Terhadap Layanan Indihome Berdasarkan Twitter Dengan Metode Klasifikasi Support Vector Machine (SVM)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 4, no. 3, p. 650, Jul. 2020, doi: 10.30865/mib.v4i3.2181.

[3] S. M. Permataning Tyas, B. S. Rintyarna, and W. Suharso, "The Impact of Feature Extraction to Naïve Bayes Based Sentiment Analysis on Review Dataset of Indihome Services," *Digital Zone: Jurnal Teknologi Informasi dan Komunikasi*, vol. 13, no. 1, pp. 1–10, Apr. 2022, doi: 10.31849/digitalzone.v13i1.9158.

[4] A. N. Nurkalyisah, A. Triayudi, and I. D. Sholihati, "Analisis Sentimen pada Twitter Berbahasa Indonesia Terhadap Penurunan Performa Layanan Indihome dan Telkomsel," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 10, no. 4, p. 387, Dec. 2022, doi: 10.26418/justin.v10i4.50858.

[5] A. Aljabar, B. M. Karomah, K. Kunci, and : Bert, "Mengungkap Opini Publik: Pendekatan BERT-based-caused untuk Analisis Sentimen pada Komentar Film," 2024.

[6] E. Yuniar and N. Hendrastuty, "Perbandingan Metode Naive Bayes, Random Forest dan SVM Untuk Analisis Sentimen Pada Twitter Tentang Kenaikan Gaji Guru," *Technology and Science (BITS)*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.6970.

[7] 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.

[8] Dhendra and V. Gayuh Utomo, "Benchmarking IndoBERT and Transformer Models for Sentiment Classification on Indonesian E-Government Service Reviews," *Jurnal Transformatika*, vol. 23, no. 1, pp. 86–95, Jul. 2025, doi: 10.26623/transformatika.v23i1.12095.

[9] Muhamad Agung Nulhakim, Yuliant Sibaroni, and Ku Muhammad Naim Ku Khalif, "Geospatial Sentiment Analysis Using Twitter Data on Natural Disasters in Indonesia with Support Vector Machine (SVM) Algorithm," *International Journal on Information and Communication Technology (IJoICT)*, vol. 10, no. 2, pp. 242–258, Jan. 2025, doi: 10.21108/ijoict.v10i2.1032.

[10] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 3, pp. 746–757, Aug. 2023, doi: 10.26555/jiteki.v9i3.26490.

[11] D. Fatma Sjoraida, B. Wibawa, K. Guna, and D. Yudhakusuma, "Analisis Sentimen Film Dirty Vote Menggunakan BERT (Bidirectional Encoder Representations from Transformers)," *Universitas Langlangbuana*, vol. 8, no. 2, 2024, doi: 10.35870/jti.

[12] A. Elhan, M. Kusuma, 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 Sentiment Analysis of Twitter Users on COVID-19 Vaccines in Indonesia using Random Forest and BERT Algorithms", [Online]. Available: https://jurnal.ipb.ac.id/index.php/jika

Vol. 6, No. 5, October 2025, Page. 2975-2990
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.4993

[13] R. Mas, R. Wahyu, P. Kusuma Atmaja, and W. Yustanti, "Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers)," *JEISBI*, vol. 02, p. 2021.

[14] Y. Roh, G. Heo, and S. E. Whang, "A Survey on Data Collection for Machine Learning: a Big Data -- AI Integration Perspective," Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.03402

[15] B. Wilie *et al.*, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding." [Online]. Available: https://github.com/annisanurulazhar/absa-playground

[16] Muhamad Agung Nulhakim, Yuliant Sibaroni, and Ku Muhammad Naim Ku Khalif, "Geospatial Sentiment Analysis Using Twitter Data on Natural Disasters in Indonesia with Support Vector Machine (SVM) Algorithm," *International Journal on Information and Communication Technology (IJoICT)*, vol. 10, no. 2, pp. 242–258, Jan. 2025, doi: 10.21108/ijoict.v10i2.1032.

[17] C. Shaw, P. LaCasse, and L. Champagne, "Exploring emotion classification of indonesian tweets using large scale transfer learning via IndoBERT," *Soc Netw Anal Min*, vol. 15, no. 1, Dec. 2025, doi: 10.1007/s13278-025-01439-6.

[18] J. Ma, W. Wang, and Q. Jin, "A Survey of Deep Learning for Geometry Problem Solving," Jul. 2025, [Online]. Available: http://arxiv.org/abs/2507.11936

[19] L. Geni, E. Yulianti, and D. I. Sensuse, "Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 9, no. 3, pp. 746–757, Aug. 2023, doi: 10.26555/jiteki.v9i3.26490.

[20] M. Agung Nulhakim, Y. Sibaroni, and K. Muhammad Naim Ku Khalif, "Geospatial Sentiment Analysis Using Twitter Data on Natural Disasters in Indonesia with Support Vector Machine," *Intl. Journal on ICT*, vol. 10, no. 2, pp. 242–258, 2024, doi: 10.21108/ijoict.v10i2.1032.

[21] H. Yan and D. Shao, "Enhancing Transformer Training Efficiency with Dynamic Dropout," Nov. 2024, [Online]. Available: http://arxiv.org/abs/2411.03236

[22] C. H. Lin and U. Nuha, "Sentiment analysis of Indonesian datasets based on a hybrid deep-learning strategy," *J Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00782-9.

[23] I. Daqiqil, H. Saputra, Syamsudhuha, R. Kurniawan, and Y. Andriyani, "Sentiment analysis of student evaluation feedback using transformer-based language models," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 2, pp. 1127–1139, Nov. 2024, doi: 10.11591/ijeecs.v36.i2.pp1127-1139.

[24] A. Jazuli, Widowati, and R. Kusumaningrum, "Optimizing Aspect-Based Sentiment Analysis Using BERT for Comprehensive Analysis of Indonesian Student Feedback," *Applied Sciences (Switzerland)*, vol. 15, no. 1, Jan. 2025, doi: 10.3390/app15010172.

[25] H. Yin, X. Song, S. Yang, and J. Li, "Sentiment analysis and topic modeling for COVID-19 vaccine discussions," *World Wide Web*, vol. 25, no. 3, pp. 1067–1083, May 2022, doi: 10.1007/s11280-022-01029-y.