Journal Collabits
Vol. 3 No. 1 January 2026 : 38-55
E-ISSN : 3046-6709, P-ISSN : 3062-8601
https://publikasi.mercubuana.ac.id/index.php/collabits

Comparative Analysis of Public Sentiment Towards Sri Mulyani and Purbaya as Finance Ministers on the X Platform Using the Indobertweet Model

Muhammad Aryaka Zamzami[1*], Siti Maesaroh[2], Dendy Jonas Managas[3]

[1,2] Informatics Engineering, Universitas Mercu Buana, Indonesia
[3] Informatics Engineering, Universitas Raharja, Indonesia

*Coressponden Author: 41522010100@student.mercubuana.ac.id

Abstract - The development of social media has positioned platform X (Twitter) as a primary source for expressing public opinion toward government figures and policies. This study aims to analyze public sentiment toward two Indonesian public figures, Sri Mulyani Indrawati and Purbaya Yudhi Sadewa, by utilizing the transformer-based IndoBERTweet model. The data were collected from January 1, 2025, to November 1, 2025. A total of 11,000 tweets related to Sri Mulyani were collected; however, only 2,500 tweets were used for data processing and model training, with a maximum limit of 1,000 tweets per month. Meanwhile, 650 tweets were obtained for Purbaya Yudhi Sadewa. This research employs a supervised learning approach with labeled data consisting of positive, negative, and neutral sentiment classes. Minimal preprocessing was applied, considering that IndoBERTweet is specifically designed to handle the characteristics of social media text. The model was trained for five epochs and evaluated using accuracy, precision, recall, and F1-score metrics. The results indicate that the IndoBERTweet model can classify sentiment effectively, particularly on the Sri Mulyani dataset, which contains a larger volume of data and achieves an accuracy of over 82%. In contrast, the model's performance on the Purbaya Yudhi Sadewa dataset shows a lower accuracy of 71%, influenced by the limited amount of data. This study confirms that the quantity and distribution of data significantly affect the performance of transformer-based sentiment analysis models. Based on the sentiment classification results, public sentiment toward Sri Mulyani Indrawati tends to be dominated by negative and neutral sentiments, while sentiment toward Purbaya Yudhi Sadewa shows a distribution dominated by neutral and positive sentiments.

Keywords :
Sentiment Analysis;
IndoBERTweet;
Social Media;
Transformer;
Public Opinion;

Article History:
Received: 28-11-2025
Revised: 16-12-2025
Accepted: 08-01-2026

Article DOI : 10.22441/collabits.v3i1.37962

INTRODUCTION

The rapid development of social media has transformed how people express opinions on public issues, including views on government policies and state officials [1]. Platform X, formerly known as Twitter, is one of the most actively used social media platforms for delivering direct and real-time public commentary, making it a valuable data source for researchers seeking to understand the dynamics of public opinion. The transition of the Minister of Finance position from Sri Mulyani Indrawati to Purbaya Yudhi Sadewa in 2025 generated diverse public responses. Given the distinct backgrounds and leadership styles of these two figures, it is reasonable to expect variations in public sentiment expressed on social media.

However, analyzing public opinion on social media cannot rely on traditional text analysis methods due to the unique linguistic characteristics of platforms such as X. Social media text is typically informal, containing abbreviations, slang, emoticons, hashtags, and distinctive conversational styles. Consequently,

natural language processing (NLP) models specifically designed to capture the nuances of Indonesian social media language are required. IndoBERTweet is one of the most relevant models for this purpose, as it is pre-trained on large-scale Indonesian Twitter data and optimized for informal text [2]. Koto et al. introduced IndoBERTweet as a Twitter-specific pre-trained language model for Indonesian and demonstrated its superior performance across various NLP tasks, including sentiment analysis [2]. This provides a strong foundation for its use in studies examining public opinion on platform X.

Furthermore, research conducted by Khairani et al. compared the application of IndoBERT and IndoBERTweet in emotion analysis of comments on Indonesian news Instagram accounts and found that IndoBERTweet outperformed IndoBERT when dealing with informal social media text [3]. These findings further support the suitability of IndoBERTweet for sentiment analysis tasks involving non-standard Indonesian language commonly found on Twitter. The relevance of this study lies in demonstrating that preprocessing strategies and model selection play a crucial role in improving the accuracy of sentiment analysis for Indonesian social media data.

Considering that the change in the Minister of Finance position directly influences public perceptions of national fiscal policy, and supported by the availability of NLP models capable of understanding social media language, a comparative analysis of public sentiment toward Sri Mulyani and Purbaya represents an important and timely research topic. This study not only provides insights into public perceptions during a leadership transition but also contributes empirical evidence on the effectiveness of IndoBERTweet for political sentiment analysis in the Indonesian context.

LITERATURE REVIEW

Based on previous studies, IndoBERTweet consistently demonstrates superior performance in sentiment analysis of Indonesian-language social media data. Most comparative research shows that IndoBERTweet outperforms models such as IndoBERT, mBERT, LSTM, SVM, FastText, and Word2Vec, particularly in handling informal, dynamic language typical of Twitter/X, including abbreviations, emojis, and colloquial expressions [4],[5],[6],[7],[8]. This advantage is largely due to IndoBERTweet being pretrained on a large-scale Indonesian Twitter corpus, enabling stronger contextual understanding of social media discourse.

Prior studies also emphasize the importance of text preprocessing in improving sentiment classification performance, although its impact may vary depending on dataset characteristics [3]. Common preprocessing steps such as removing URLs and mentions, as well as normalizing non-standard words, generally contribute to better model performance. In parallel, advances in Transformer-based architectures have significantly improved contextual sensitivity and classification accuracy, allowing models like IndoBERTweet to consistently outperform traditional machine learning methods and classical word embedding approaches.

IndoBERTweet has been successfully applied across diverse domains, including economics, politics, public policy, entertainment, and hate speech detection, with stable and reliable evaluation results [9],[10]. Some studies further report performance improvements when IndoBERTweet is combined with architectures such as BiLSTM, achieving accuracy and F1-scores above 93% [10]. Overall, existing literature provides strong evidence that IndoBERTweet is well suited for sentiment analysis of public conversations on platform X, particularly given the linguistic complexity of Indonesian social media content.

Despite the growing body of research on sentiment analysis and Transformer-based models in Indonesia, several gaps remain. First, studies that explicitly compare public sentiment toward two Ministers of Finance during a leadership transition, Sri Mulyani Indrawati and Purbaya Yudhi Sadewa are still scarce, even though such transitions may rapidly influence public perception of fiscal policy.

Second, while IndoBERTweet has demonstrated strong performance in general Twitter-based tasks, focused evaluations within political–economic discourse and leadership transition contexts remain limited. Most existing studies concentrate on general topics, hate speech detection, or non-political domains.

Third, many sentiment analysis studies remain fragmented, emphasizing classification results without integrating topic or keyword analysis to explain the underlying drivers of sentiment. Combining sentiment classification with keyword extraction techniques can provide deeper insights into dominant public narratives.

Finally, research using post-2021 X (Twitter) data to analyze temporal sentiment dynamics during official transitions is still limited. Therefore, empirical studies covering the 2024–2025 period are needed to update and strengthen existing evidence.

Accordingly, this study aims to address these gaps by conducting a comparative sentiment analysis of public opinions toward Sri Mulyani Indrawati and Purbaya Yudhi Sadewa on platform X, while evaluating the performance of the IndoBERTweet model within a political–economic context.

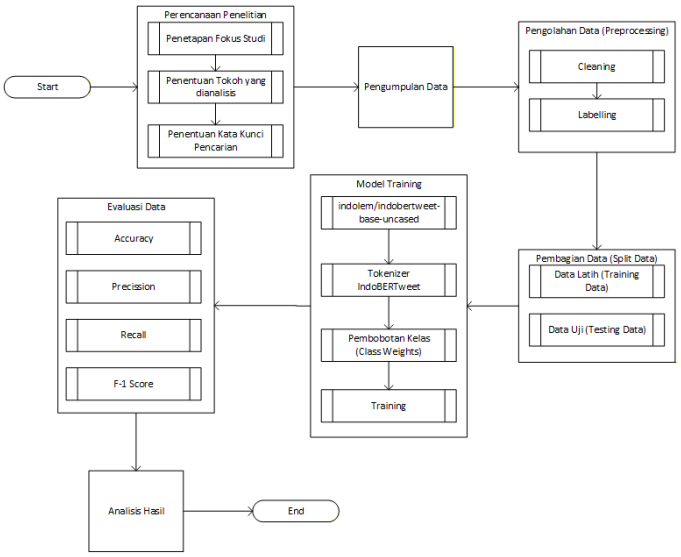
METHODOLOGY

This study adopts a quantitative research approach using computational analysis based on Natural Language Processing (NLP) to measure and compare public sentiment toward Sri Mulyani Indrawati and Purbaya Yudhi Sadewa on platform X (Twitter). A quantitative approach is selected because sentiment analysis involves processing large volumes of textual data that are subsequently transformed into numerical values or categorical labels (positive, negative, and neutral). This transformation enables objective measurement, comparison, and evaluation of model performance.

This approach aligns with the fundamental characteristics of sentiment analysis research, which commonly emphasizes machine learning–based text classification. As noted by Alhaq et al., sentiment analysis is inherently a quantitative methodology that relies on NLP and machine learning algorithms to model and classify textual data [11]. Therefore, the use of a quantitative computational approach is considered appropriate for systematically analyzing public sentiment expressed on social media platforms.

3.1 Research Stages

Figure 1. Reasearch Stages



The first stage of the research is the planning phase, which includes defining the research focus, determining the public figures to be analyzed, and selecting relevant search keywords. At this stage, the research objectives are formulated, and the scraping time frame is established to capture the dynamics of public opinion toward Sri Mulyani Indrawati and Purbaya Yudhi Sadewa. The technical instruments to be used, namely Tweet-Harvest and Python scripts are determined to ensure that the entire data collection process is conducted in an automated and consistent manner.

The second stage is data collection, which involves scraping tweets from platform X using Tweet-

Harvest with predefined parameters, including keywords, date range, and Indonesian language settings. The scraping process is conducted incrementally using monthly time intervals to avoid technical constraints such as rate limits. Each scraping result is stored in a separate CSV file and subsequently merged using Python scripts to produce a single comprehensive dataset. During this stage, duplicate tweets are removed, and the data are sorted chronologically to ensure dataset integrity.

The third stage is data preprocessing, which focuses on cleaning the textual data. This process includes removing URLs, punctuation marks, mentions, hashtags, numbers, emojis, and other irrelevant textual elements that may introduce noise and affect model performance.

The fourth stage is data splitting. The dataset is divided into two subsets: training data and testing data. This division is performed using a stratified split method to maintain proportional sentiment class distributions in both subsets.

The fifth stage involves sentiment modeling using the IndoBERTweet model. The process begins with case folding using the indolem/indobertweet-base-uncased model, followed by tokenization performed by the IndoBERTweet tokenizer. To address class imbalance in sentiment distribution, class weighting is applied prior to the training process. The final step of this stage is model training using the prepared training dataset.

After the training process is completed, the sixth stage is model evaluation. Model performance is assessed using accuracy, precision, recall, and F1-score as the primary evaluation metrics. These metrics are employed to ensure that the model demonstrates adequate predictive reliability before being applied to the full dataset.

The seventh stage is result interpretation. The trained model is applied to the entire collection of tweets to generate final sentiment classifications. The resulting sentiment data are further analyzed to obtain insights such as sentiment distribution for each public figure, comparative trends in public opinion, and sentiment trend visualizations in graphical form. Additionally, supplementary aappears conducted by identifying dominant keywords frequently appearing in tweets to gain a deeper contextual understanding of public opinion.

RESULTS AND DISCUSSION

4.1 Overview of Research Data (Raw Data)

The data used in this study were obtained from platform X (Twitter) through a web scraping process using the Tweet-Harvest tool. Data collection was conducted based on keywords representing the two public figures analyzed in this research, namely “Sri Mulyani” and “Purbaya”. The data were collected over a period ranging from January 1, 2025, to November 1, 2025. This time frame was selected to capture the dynamics of public opinion both before and after the transition of the Indonesian Minister of Finance position, allowing the dataset to represent public perceptions during the leadership transition period.

Based on the data collection results, a significant disparity was observed in the number of tweets associated with each public figure. Approximately 11,000 tweets were collected for Sri Mulyani Indrawati, while around 600 tweets were obtained for Purbaya Yudhi Sadewa. This discrepancy reflects the substantially higher level of public attention toward Sri Mulyani, which may be influenced by factors such as popularity, length of tenure, and media exposure associated with each figure.

Given the large volume of tweets related to Sri Mulyani, this study applied a data limitation strategy by selecting a maximum of 1,000 tweets per month. This limitation was implemented to maintain a balanced temporal distribution of data and to prevent overrepresentation from specific months that could introduce bias into the sentiment analysis. Through this approach, the Sri Mulyani dataset remains representative of temporal opinion dynamics while ensuring computational efficiency.

Figure 2. Reasearch Stages

created_at	full_text	id_str
Fri Sep 26 14:46:51 +0000 2025	Minta tolong kasih paham ke pak dadan itu yak pak purbaya. Itu orang emg g jelas	1971587275185369278
	@ysznprd @ChatibBasri @KemenkeuRI Ternyata bener dong emang aksi sepihak	
Fri Sep 26 14:41:00 +0000 2025	Himbara. Apa karena direksi Himbara pada dendam kali ya gegara disindir Menkeu	1971585803395731926
Fri Sep 26 14:40:34 +0000 2025	Purbaya vs Rocky Gerung	1971585694327050634
Fri Sep 26 14:30:55 +0000 2025	@PDI_Perjuangan Ibuk megawati ibuk puan maharani terhormat suruh Jaksa Agung	1971583267292995999
Fri Sep 26 14:29:21 +0000 2025	seandainya semua menteri kayak pak Purbaya menarik juga ya? koboi transparan dan	1971582869840834846
	8. Buat saya stempel FIRAUN gak tepat bahkan brutal pak purbaya @KemenkeuRI ! Tapi	
Fri Sep 26 14:21:40 +0000 2025	buat pimred @mnctrijaya @gaibfiles @SINDOnews pastilah menyenangkan karena jadi	1971580939466936553
	@CNNIndonesia Kalau cuma 6% tinggal panggil BPS untuk utak-atik data (difabrikasi)	
Fri Sep 26 14:19:06 +0000 2025	terus diumumkan bahwa ekonomi tumbuh 6%. Sekarang rumusannya diganti yaitu:	1971580291698540630
Fri Sep 26 14:16:29 +0000 2025	@cnbcindonesia Ya kata Purbaya orang kaya Indonesia ramai beli dolar	1971579635508097061
Fri Sep 26 14:16:24 +0000 2025	Menkeu Purbaya jawab kritikan Rocky Gerung. https://t.co/kwnD2lY5TL	1971579614234558646
Fri Sep 26 14:12:07 +0000 2025	Utas Heran 1 Dana Pemda ngendap. Tahun ini tembus Rp233 triliun. Duit segede itu	1971578535975493706
	Bonus 10.000 saldo kalauu install apk Shopeepay Caranya : Upgrade ke Spay plus log in	
Fri Sep 26 14:11:00 +0000 2025	dg akun shopee & masukkan kode referral FF6JWV5WP diskon promo info	1971578252951994669
Fri Sep 26 14:08:20 +0000 2025	@kumparan Kayaknya pak purbaya ini the new Jokowi deh ... kayaknya cocok untuk	1971577581716836420
Fri Sep 26 13:55:14 +0000 2025	September S nya apa? Klo kata pak Purbaya S nya itu Sidak	1971574286705217610

All data obtained during the initial collection stage consisted of raw data containing various non-textual elements and noise, such as URLs, mentions, hashtags, emoticons, special characters, and potential duplicate tweets. These characteristics reflect the concise, informal, and symbol-rich communication style commonly found on platform X. Consequently, the raw data could not be directly utilized for modeling purposes and required preprocessing to improve data quality prior to further analysis.

4.2 Preprocessing Text

The data preprocessing stage in this study aims to clean tweet texts by removing various non-informative elements that may interfere with the sentiment analysis process. Preprocessing was conducted using Python scripts based on the pandas library, regular expressions, emoji handling libraries, and HTML parsers, specifically designed to address the linguistic and structural characteristics of social media data from platform X (Twitter).

The preprocessing process begins with handling missing values in the tweet text column (full_text). Tweets that do not contain textual content are removed from the dataset to prevent potential disruption in subsequent processing stages. Following this step, text cleaning is applied to each tweet using a clean_text function, which consists of several key procedures.

The text cleaning steps include decoding HTML entities, such as converting & into &, to ensure that the text is displayed and interpreted correctly as originally intended. Next, URLs whether starting with http, https, or www are removed, as they do not carry sentiment-related information. User mentions (e.g., @username) are also eliminated because they do not contribute meaningful semantic content in the context of sentiment analysis. Hashtag symbols (#) are removed while retaining the associated keywords, thereby preserving topical information. Retweet markers (RT) are excluded to minimize the influence of duplicated content.

Furthermore, all emojis are removed from the text. This decision is based on the focus of the study on text-based sentiment analysis and the capability of the IndoBERTweet model to perform effectively without explicit emoji representations. Additional cleaning is performed to remove excessive whitespace and special Unicode characters, such as zero-width spaces and non-breaking spaces, which frequently appear as artifacts of the scraping process. This step is crucial to prevent tokenization errors and to ensure consistency in text formatting.

Finally, redundant spaces are normalized, and trimming is applied to remove leading and trailing whitespace, resulting in a clean and standardized textual dataset ready for downstream modeling and sentiment classification.

Figure 3. Text Preprocessing Results

id_str	cleaned_text	full_text
1884989588391051618	Sri Mulyani tambah anggaran program MBG Rp.100 T ? Halah tagihan vendor aja belum pada dibayar.	@TirtoID Sri Mulyani tambah anggaran program MBG Rp.100 T ? Halah tagihan vendor aja belum pada dibayar.
1884988776046637331	Pak kami rakyat beli tabung elpiji 3 kg harganya di atas Rp 20 000 per tabung harga aslinya Rp 12 750 per tabung menurut ibu Sri mulyani	@kompascom Pak @prabowo kami rakyat beli tabung elpiji 3 kg harganya di atas Rp 20 000 per tabung harga aslinya Rp 12 750 per tabung menurut ibu Sri mulyani @ListyoSigitP @ahriesonta
1884984475203019229	Menteri Keuangan Sri Mulyani Indrawati mengungkapkan kekagetan saat mengetahui harga elpiji 3 kilogram (kg) di	Menteri Keuangan Sri Mulyani Indrawati mengungkapkan kekagetan saat mengetahui harga elpiji 3 kilogram (kg) di
1884981018144444809	Sri mulyani goblok	Sri mulyani goblok
1884980562819264975	Tariknya Sri Mulyani pura-pura kaget padahal harga jual di Masyarakat kami selalu 25.000 kalo susah 30.000.-	@mardigu024 Tariknya Sri Mulyani pura-pura kaget padahal harga jual di Masyarakat kami selalu 25.000 kalo susah 30.000.-

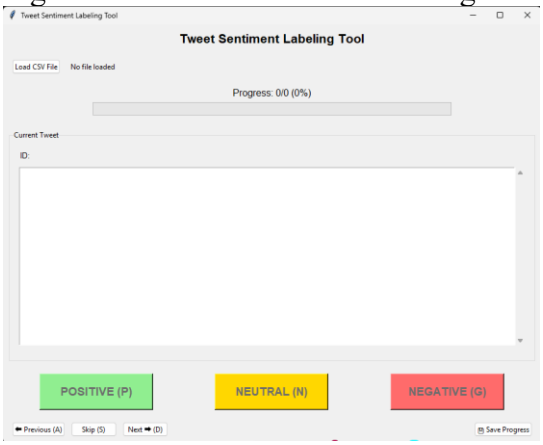
4.3 Manual Data Labeling

The data labeling stage is a crucial part of this study, as it provides sentiment labels that serve as ground truth for training and evaluating the IndoBERTweet model. Manual labeling is employed to ensure label accuracy, given the contextual, ambiguous, and implicit nature of social media language, which is often difficult to capture through automated labeling methods.

Sentiment annotation is performed on preprocessed tweet texts that have been cleaned of non-informative elements such as URLs, mentions, emojis, and special characters. Each tweet is classified into one of three sentiment categories: positive, neutral, or negative based on the overall opinion expressed toward the public figure being discussed.

To support a consistent and efficient labeling process, a custom Python-based Graphical User Interface (GUI) application is developed using the Tkinter library. The application displays tweets individually, allowing the annotator to assess each tweet's context before assigning a sentiment label. It provides three sentiment selection buttons (positive, neutral, and negative), navigation controls, and a progress bar to monitor labeling progress in real time. Labeled data can be saved periodically in CSV format to prevent data loss.

Figure 4. Tweet Sentiment Labelling GUI



Label assignment is based on the overall semantic meaning of each tweet rather than isolated keywords. Tweets expressing support or favorable evaluations are labeled as positive, those containing criticism or dissatisfaction are labeled as negative, and tweets that are informational or lack clear opinion polarity are labeled as neutral. The resulting labeled dataset is subsequently used for training and testing the IndoBERTweet model in the next stage of the study.

Figure 5. Purbaya Labeling Results

id_str	cleaned_text	sentiment
197142786330080102	Good idea Pak! Walaupun ada resikonya. Fixed Lord Purbaya tipe risk taker.	positive
198439764805774986	Kalau jadi Purbaya lebih baik mundur sudah tidak di anggap ngapain lagi	neutral
194107706139077433	Bursa Komisioner LPS 2025-2030: Purbaya Kembali Ajukan Diri	neutral
189119353897896763	Sadranan Agung Wotgaleh merupakan sebuah tradisi tahunan yang diselenggarakan di kompleks Masjid Sulthoni Wot Galeh Sendangtirto Berbah Sleman sebuah bangunan bersejarah yang menjadi	neutral
193557262365381037	pasti isi kepalamu bukan otak tapi serpihan cendol	negative
197195819853937493	Beneran gak nih menkeu Purbaya Jangan kayak wakil BGN Nipu nangis ye Males	neutral
190536675581779608	Situasi yang selalu di rindukan saudara perantau saat Mudik kali ini saya bersama dengan beberapa temen nyambangin terminal Purbaya Surabaya Jawa Timur senengnya berbagi kebahagiaan Selamat Mudik ya dulur hati-hati di jalan ya mari ciptakan Mudik aman dan Nyaman	neutral
197183966741184526	RG tidak harus minta maaf! justru kritikan RG menjadi cambuk bagi Purbaya utk membuktikan kemampuannya.	positive

4.4 Dividing Training Data and Test Data

After all tweet data had undergone preprocessing and manual sentiment labeling, the next step was to split the dataset into training and testing sets. This data partitioning aims to evaluate the generalization capability of the IndoBERTweet model in classifying sentiment on previously unseen data.

Prior to data splitting, only tweets with valid sentiment labels (positive, neutral, and negative) were included in the analysis. Tweets with missing or undefined labels were excluded to ensure the quality of both training and testing datasets. The labeled dataset was then examined through Exploratory Data

Analysis (EDA) to understand sentiment class distributions and text length characteristics across classes.

Figure 6. Sri Mulyani Sentiment Distribution Results

```
================================================
EXPLORATORY DATA ANALYSIS
================================================

1. Sentiment Distribution:
sentiment
negative    1014
neutral      934
positive     871
Name: count, dtype: int64

Percentages:
sentiment
negative    35.970202
neutral     33.132316
positive    30.897481
Name: proportion, dtype: float64

Class Imbalance Ratio: 1.16
```

Figure 7. Sri Mulyani Sentiment Distribution Chart Results

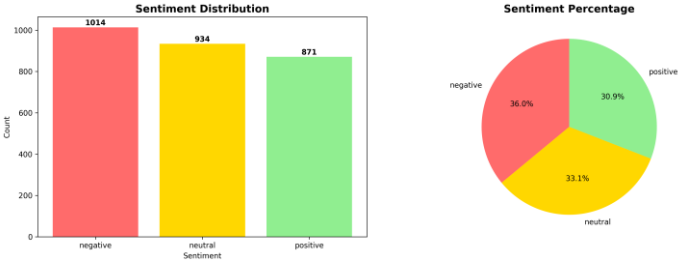
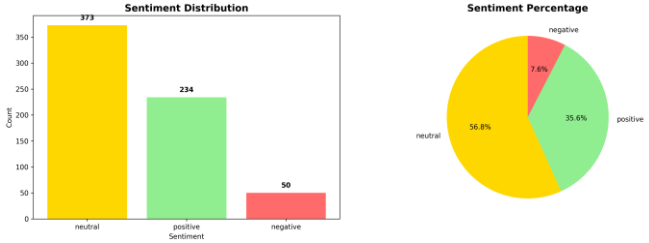


Figure 8. Purbaya Sentiment Distribution Results

```
================================================
EXPLORATORY DATA ANALYSIS
================================================

1. Sentiment Distribution:
sentiment
neutral     373
positive    234
negative     50
Name: count, dtype: int64

Percentages:
sentiment
neutral     56.773212
positive    35.616438
negative     7.610350
Name: proportion, dtype: float64

Class Imbalance Ratio: 7.46
WARNING: Significant class imbalance detected!
```

Figure 9. Purbaya Sentiment Distribution Chart Results



The sentiment distribution analysis revealed class imbalance, particularly in the dataset related to Minister of Finance Purbaya, indicating unequal representation among sentiment classes. To address this issue, a stratified split method was applied, ensuring that the proportion of each sentiment class was

preserved in both the training and testing sets. This approach helps prevent class dominance in either subset, which could otherwise bias model performance and evaluation [9].

In this study, the dataset was divided using an 80:20 ratio for training and testing data, respectively. The splitting process was performed using the *train_test_split* function from the scikit-learn library with the stratify parameter enabled, while a fixed random_state was applied to ensure reproducibility.

Figure 10. Purbaya Data Split Results

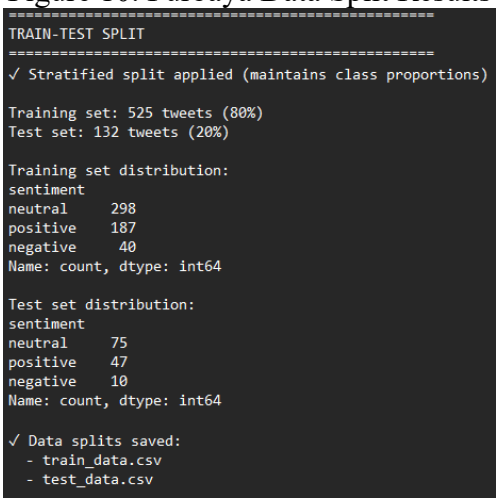
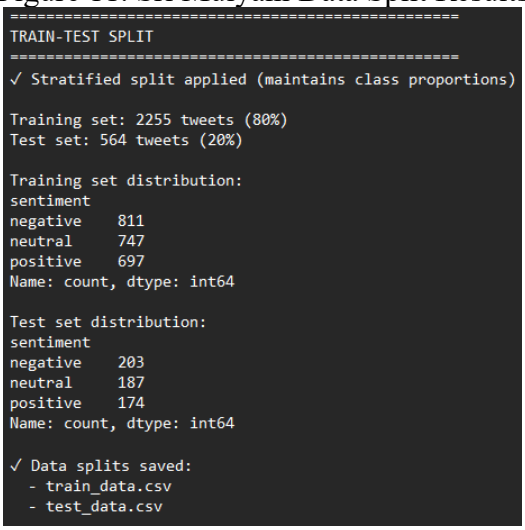


Figure 11. Sri Mulyani Data Split Results



The resulting training data were used to train the IndoBERTweet model, whereas the testing data were utilized to evaluate model performance using accuracy, precision, recall, and F1-score metrics. Both subsets were stored as separate CSV files to ensure consistency in subsequent modeling and evaluation stages.

4.5 IndoBERTweet Model Training

a. Initializing the Model and Tokenizer

The model employed in this study is indolem/indobertweet-base-uncased. Both the tokenizer and the model are loaded using the Hugging Face Transformers library. The tokenizer converts raw text into numerical representations in the form of token IDs that can be processed by the model, while the model performs sentiment classification into three classes: negative, neutral, and positive.

Figure 12. Library in training.py

```
import pandas as pd
import numpy as np
import torch
from torch.utils.data import Dataset, DataLoader
from transformers import (
    AutoTokenizer,
    AutoModelForSequenceClassification,
    TrainingArguments,
    Trainer,
    EarlyStoppingCallback
)
from sklearn.metrics import accuracy_score, precision_recall_fscore_support,
classification_report, confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.utils.class_weight import compute_class_weight
```

Each sentiment label is mapped to a numerical value, with negative labeled as 0, neutral as 1, and positive as 2. This encoding is required for training the deep learning–based model. The training process is executed on the available hardware, utilizing either a GPU or CPU to optimize computational efficiency.

b. Tokenization

The training and testing data obtained from the previous stage are converted into a PyTorch *Dataset* format using the *SentimentDataset* class. At this stage, each tweet text is processed with the IndoBERTweet tokenizer using a maximum sequence length of 128 tokens. Texts exceeding this limit are truncated, while shorter sequences are padded to ensure uniform input length.

In addition, the tokenizer generates attention masks that enable the model to distinguish between actual tokens and padding tokens. This tokenization and dataset construction process ensures that all data are properly formatted and ready for use in the model training and evaluation stages.

Figure 13. Tokenization

```
class SentimentDataset(Dataset):
    """Custom Dataset for sentiment analysis"""
    def __init__(self, texts, labels, tokenizer, max_length=128):
        self.texts = texts
        self.labels = labels
        self.tokenizer = tokenizer
        self.max_length = max_length

    def __len__(self):
        return len(self.texts)

    def __getitem__(self, idx):
        text = str(self.texts.iloc[idx])
        label = self.labels.iloc[idx]

        encoding = self.tokenizer(
            text,
            add_special_tokens=True,
            max_length=self.max_length,
            padding='max_length',
            truncation=True,
            return_attention_mask=True,
            return_tensors='pt'
        )

        return {
            'input_ids': encoding['input_ids'].flatten(),
            'attention_mask': encoding['attention_mask'].flatten(),
            'labels': torch.tensor(label, dtype=torch.long)
        }
```

c. Handling Class Imbalance

Based on the sentiment distribution analysis of the training data, class imbalance was observed among sentiment categories, particularly in the dataset related to Purbaya. To address this issue, a class weighting approach was applied, with class weights computed using the balanced class weight method from the scikit-learn library.

Figure 14. Class Weighting

```
class IndoBERTweetTrainer:

    def compute_class_weights(self):
        """Compute class weights to handle imbalance"""
        labels = self.train_df['label_id'].values
        class_weights = compute_class_weight(
            class_weight='balanced',
            classes=np.unique(labels),
            y=labels
        )
        print(f"\n Class weights (to handle imbalance):")
        for i, weight in enumerate(class_weights):
            print(f"  {self.id2label[i]}: {weight:.2f}")

        return torch.tensor(class_weights, dtype=torch.float).to(self.device)
```

These class weights were incorporated into the CrossEntropyLoss function, enabling the model to assign greater importance to underrepresented sentiment classes. This strategy aims to improve the model's ability to learn from all sentiment categories more evenly and to reduce bias toward the majority class.

Figure 15. CrossEntropyLoss

```
class IndoBERTweetTrainer:

    def train(self, output_dir='./indobertweet_sentiment_model_mulyani',
              epochs=5, batch_size=16, learning_rate=2e-5,
              use_class_weights=True):
        """Train the model"""
        print("\n" + "="*50)
        print("TRAINING MODEL")
        print("="*50)

        # Prepare datasets
        train_dataset, test_dataset = self.prepare_datasets()

        # Compute class weights
        if use_class_weights:
            class_weights = self.compute_class_weights()

            # Custom Trainer with class weights
            class WeightedTrainer(Trainer):
                def compute_loss(self, model, inputs, return_outputs=False,
                num_items_in_batch=None):
                    labels = inputs.pop("labels")
                    outputs = model(**inputs)
                    logits = outputs.logits
                    loss_fct = torch.nn.CrossEntropyLoss(weight=class_weights)
                    loss = loss_fct(logits, labels)
                    return (loss, outputs) if return_outputs else loss

            trainer_class = WeightedTrainer
        else:
            trainer_class = Trainer
```

d. Model Training Configuration

Model training is conducted using the Trainer class from the Transformers library with the following configurations: (a) five training epochs, (b) a batch size of 16 for both training and testing data, (c) a learning rate of 2×10^{-5}, (d) the AdamW optimizer with a weight decay of 0.01, and (e) 100 warm-up steps.

Model evaluation is performed at the end of each epoch using the testing dataset (evaluation strategy: epoch). The best-performing model is selected based on the highest F1-score, as this metric is considered most appropriate for classification tasks involving class imbalance. In addition, an early stopping mechanism with a patience of three epochs is applied to prevent overfitting.

Figure 16. Model Training Configuration

```python
class IndoBERTweetTrainer:
    def train(self, output_dir='./indobertweet_sentiment_model_mulyani',
        # Training arguments
        training_args = TrainingArguments(
            output_dir=output_dir,
            num_train_epochs=epochs,
            per_device_train_batch_size=batch_size,
            per_device_eval_batch_size=batch_size,
            learning_rate=learning_rate,
            warmup_steps=100,
            weight_decay=0.01,
            logging_dir='./logs',
            logging_steps=10,
            eval_strategy="epoch",
            save_strategy="epoch",
            load_best_model_at_end=True,
            metric_for_best_model="f1",
            greater_is_better=True,
            save_total_limit=2,
            report_to="none"
        )

        # Initialize trainer
        self.trainer = trainer_class(
            model=self.model,
            args=training_args,
            train_dataset=train_dataset,
            eval_dataset=test_dataset,
            compute_metrics=self.compute_metrics,
            callbacks=[EarlyStoppingCallback(early_stopping_patience=3)]
        )
```

e. Model Training Process

Model training is performed by iteratively feeding the training data into the IndoBERTweet model across the predefined number of epochs. During this process, the model continuously updates its internal parameters to minimize the loss function and improve sentiment classification performance.

After training is completed, the best-performing model is automatically reloaded and saved for evaluation. The trained model and its corresponding tokenizer are then stored in a designated directory for subsequent testing and potential deployment.

Figure 17. Results of the Purbaya's Model Training

```
{'loss': 0.9095, 'grad_norm': 129.6024169921875, 'learning_rate': 1.38e-05, 'epoch': 2.12}
{'loss': 0.8464, 'grad_norm': 7.982986927032471, 'learning_rate': 1.58e-05, 'epoch': 2.42}
{'loss': 0.77, 'grad_norm': 5.25533390045166, 'learning_rate': 1.7800000000000002e-05, 'epoch': 2.73}
{'eval_loss': 0.7576563358306885, 'eval_accuracy': 0.6590909090909091, 'eval_f1': 0.6754163124901762,
'eval_precision': 0.7375056809019073, 'eval_recall': 0.6590909090909091, 'eval_runtime': 21.5529,
'eval_samples_per_second': 6.124, 'eval_steps_per_second': 0.418, 'epoch': 3.0}
 60%|                                                                                | 99/165
[17:53<10:40,  9.71s/it]C:\Users\aryaka\AppData\Local\Programs\Python\Python310\lib\site-packages\torch
\utils\data\dataloader.py:665: UserWarning: 'pin_memory' argument is set as true but no accelerator is
found, then device pinned memory won't be used.
  warnings.warn(warn_msg)
{'loss': 0.7432, 'grad_norm': 6.114866733551025, 'learning_rate': 1.98e-05, 'epoch': 3.03}
{'loss': 0.6479, 'grad_norm': 5.738476753234863, 'learning_rate': 1.7230769230769234e-05, 'epoch': 3.33}
{'loss': 0.6266, 'grad_norm': 6.029093265533447, 'learning_rate': 1.4153846153846156e-05, 'epoch': 3.64}
{'loss': 0.5383, 'grad_norm': 6.454534053802249, 'learning_rate': 1.1076923076923079e-05, 'epoch': 3.94}
{'eval_loss': 0.7797732353210449, 'eval_accuracy': 0.7121212121212122, 'eval_f1': 0.7121212121212122,
'eval_precision': 0.7121212121212122, 'eval_recall': 0.7121212121212122, 'eval_runtime': 21.4431,
'eval_samples_per_second': 6.156, 'eval_steps_per_second': 0.42, 'epoch': 4.0}
 80%|                                                                                | 132/165
[23:49<05:24,  9.84s/it]C:\Users\aryaka\AppData\Local\Programs\Python\Python310\lib\site-packages\torch
\utils\data\dataloader.py:665: UserWarning: 'pin_memory' argument is set as true but no accelerator is
found, then device pinned memory won't be used.
  warnings.warn(warn_msg)
{'loss': 0.5193, 'grad_norm': 7.1074347496032715, 'learning_rate': 8.000000000000001e-06, 'epoch': 4.24}
{'loss': 0.4819, 'grad_norm': 11.518919944763184, 'learning_rate': 4.923076923076924e-06, 'epoch': 4.55}
{'loss': 0.4518, 'grad_norm': 7.685436725616455, 'learning_rate': 1.8461538461538465e-06, 'epoch': 4.85}
{'eval_loss': 0.7382343411445618, 'eval_accuracy': 0.7196969696969697, 'eval_f1': 0.7218855218855219,
'eval_precision': 0.7271468841236284, 'eval_recall': 0.7196969696969697, 'eval_runtime': 15.8,
'eval_samples_per_second': 8.354, 'eval_steps_per_second': 0.57, 'epoch': 5.0}
{'train_runtime': 1771.0265, 'train_samples_per_second': 1.482, 'train_steps_per_second': 0.093,
'train_loss': 0.7981757915381229, 'epoch': 5.0}
100%|                                                                                |
165/165 [29:31<00:00, 10.73s/it]
```

Figure 18. Results of the Sri Mulyani's Model Training

```
{'loss': 0.1454, 'grad_norm': 5.745270252227783, 'learning_rate': 4.495867768595042e-06, 'epoch': 4.04}
{'loss': 0.0845, 'grad_norm': 6.812795162200928, 'learning_rate': 4.165289256198348e-06, 'epoch': 4.11}
{'loss': 0.1094, 'grad_norm': 4.5062026977539006, 'learning_rate': 3.834710743801654e-06, 'epoch': 4.18}
{'loss': 0.0574, 'grad_norm': 1.6863874197006226, 'learning_rate': 3.504132231404959e-06, 'epoch': 4.26}
{'loss': 0.1018, 'grad_norm': 9.376374244689941, 'learning_rate': 3.1735537190082656e-06, 'epoch': 4.33}
{'loss': 0.092, 'grad_norm': 4.814056873321533, 'learning_rate': 2.8429752066115707e-06, 'epoch': 4.4}
{'loss': 0.09, 'grad_norm': 13.258831977844238, 'learning_rate': 2.5123966942148765e-06, 'epoch': 4.47}
{'loss': 0.0755, 'grad_norm': 10.482351303100586, 'learning_rate': 2.1818181818181824e-06, 'epoch': 4.54}
{'loss': 0.089, 'grad_norm': 5.574278831481934, 'learning_rate': 1.8512396694214876e-06, 'epoch': 4.61}
{'loss': 0.1365, 'grad_norm': 2.0958771705627441, 'learning_rate': 1.5206611570247934e-06, 'epoch': 4.68}
{'loss': 0.1075, 'grad_norm': 8.28422737121582, 'learning_rate': 1.1900826446280993e-06, 'epoch': 4.75}
{'loss': 0.1623, 'grad_norm': 13.323905944824219, 'learning_rate': 8.595041322314054e-07, 'epoch': 4.82}
{'loss': 0.0831, 'grad_norm': 3.961151361465454, 'learning_rate': 5.289256198347107e-07, 'epoch': 4.89}
{'loss': 0.1264, 'grad_norm': 3.8718347549438477, 'learning_rate': 1.9834710743801655e-07, 'epoch': 4.96}
{'eval_loss': 0.5331471562385559, 'eval_accuracy': 0.8138297872340425, 'eval_f1': 0.8165404642486024,
'eval_precision': 0.8233986115967684, 'eval_recall': 0.8138297872340425, 'eval_runtime': 88.8128,
'eval_samples_per_second': 6.35, 'eval_steps_per_second': 0.405, 'epoch': 5.0}
{'train_runtime': 7517.3765, 'train_samples_per_second': 1.5, 'train_steps_per_second': 0.094,
'train_loss': 0.34597235816590327, 'epoch': 5.0}
100%|                                                                                |
705/705 [2:05:17<00:00, 10.66s/it]
```

In the case of Sri Mulyani, the training process required approximately ±2 hours to complete five epochs. The training results show a consistent decrease in loss values, from approximately 1.1 at the beginning of training to 0.1 at the final epoch, indicating that the model successfully learned patterns from the training data. In addition, accuracy and F1-score remained stable above 81%, suggesting that the model achieved satisfactory convergence without significant overfitting.

In contrast, for Purbaya Yudhi Sadewa, the smaller dataset resulted in a shorter training time of approximately ±30 minutes for five epochs. The training results exhibit a similar trend, with the loss decreasing from approximately 1.08 to 0.45 and accuracy improving from 53% to around 71%. The F1-score also increased to approximately 72%, indicating that the model was still able to learn sentiment patterns effectively despite the relatively limited amount of data.

4.6 Model Evaluation and Result Analysis

The model evaluation stage is conducted to assess the performance of the IndoBERTweet model in classifying tweet sentiments into three classes: negative, neutral, and positive. Evaluation is performed using the testing dataset separated in the previous stage. The evaluation methods include a confusion matrix and standard performance metrics, namely accuracy, precision, recall, and F1-score. Model evaluation is carried out separately for the two research subjects, namely public sentiment toward Sri Mulyani Indrawati and Purbaya Yudhi Sadewa.

a. Model Evaluation on Sri Mulyani Indrawati

Based on the evaluation results on the Sri Mulyani Indrawati sentiment testing dataset, the IndoBERTweet model achieved an accuracy of 82.45%, with an F1-score of 0.8269, precision of 0.8328, and recall of 0.8245. These results indicate that the model demonstrates strong overall performance in sentiment classification.

Table 1. Overall Metrics Sri Mulyani

Category	Overall Metrics
Accuracy	0.8245
F1-Score	0.8269
Precision	0.8328
Recall	0.8245

According to the classification report, the neutral sentiment class achieved the best performance, with a precision of 0.97 and a recall of 0.85, indicating that most neutral tweets were correctly identified. The negative and positive sentiment classes also exhibited balanced performance, with F1-scores of 0.81 and 0.77, respectively (Figure 4.31).

Figure 19. Classification Report Sri Mulyani

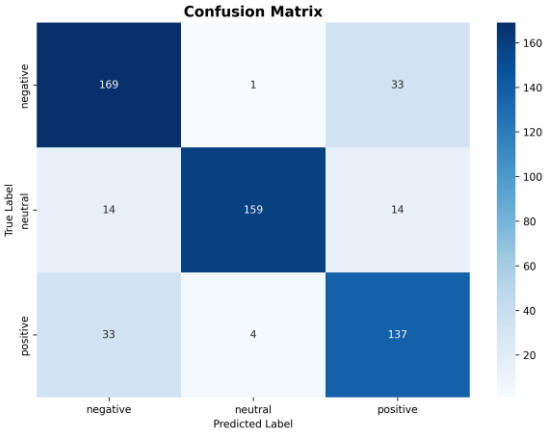
Classification Report:

	precision	recall	f1-score	support
negative	0.78	0.83	0.81	203
neutral	0.97	0.85	0.91	187
positive	0.74	0.79	0.77	174
accuracy			0.82	564
macro avg	0.83	0.82	0.83	564
weighted avg	0.83	0.82	0.83	564

Further analysis using the confusion matrix shows that the majority of tweets in each sentiment class were correctly classified, as reflected by high values along the main diagonal. However, misclassifications were still observed primarily between the negative and positive classes. This suggests that the model occasionally struggles to distinguish between critical yet constructive opinions and seemingly appreciative statements that convey opposite meanings, such as sarcasm, particularly in tweets discussing

complex economic policies.

Figure 20. Confusion Matrix Sri Mulyani



Overall, these evaluation results demonstrate that the IndoBERTweet model is effective in capturing public sentiment patterns toward Sri Mulyani Indrawati, supported by a relatively large training dataset and a more balanced sentiment class distribution.

b. Model Evaluation on Purbaya

The evaluation results on the Purbaya Yudhi Sadewa sentiment dataset indicate lower performance compared to the results obtained for Sri Mulyani Indrawati. The IndoBERTweet model achieved an accuracy of 71.97%, with an F1-score of 0.7219, precision of 0.7271, and recall of 0.7197.

Table 2. Overall Metrics Purbaya

Category	Overall Metrics
Accuracy	0.7197
F1-Score	0.7219
Precision	0.7271
Recall	0.7197

According to the classification report, the neutral sentiment class again demonstrated the strongest performance, achieving an F1-score of 0.80. This suggests that the model remains relatively stable in identifying tweets that are primarily informational or descriptive. In contrast, the negative sentiment class exhibited the weakest performance, with a precision of 0.43 and a recall of 0.60, indicating that a substantial portion of negative tweets were misclassified as neutral or positive.

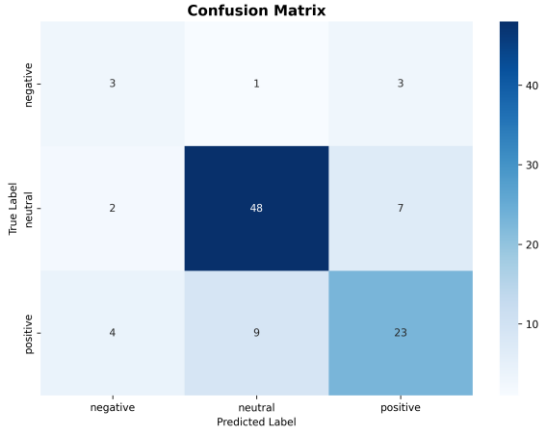
Figure 21. Classification Report Purbaya

```
Classification Report:
              precision    recall  f1-score   support

    negative       0.43      0.60      0.50        10
     neutral       0.80      0.80      0.80        75
    positive       0.67      0.62      0.64        47

    accuracy                           0.72       132
   macro avg       0.63      0.67      0.65       132
weighted avg       0.73      0.72      0.72       132
```

The confusion matrix for the Purbaya Yudhi Sadewa dataset reveals frequent misclassifications between the negative and positive sentiment classes. This outcome is largely influenced by the relatively small size of the testing dataset, particularly for the negative class, which contains only a limited number of samples. The imbalance in class distribution reduces the model's ability to learn and represent negative

sentiment characteristics effectively.

Figure 22. Confusion Matrix Purbaya



c. Word Cloud Analysis and Sentiment Context Interpretation

Word cloud visualization is employed in this study as an exploratory method to identify dominant keywords and topics frequently appearing in public discussions related to Sri Mulyani Indrawati and Purbaya Yudhi Sadewa on platform X. The word clouds are generated based on word frequency from tweet data that have undergone preprocessing. Words that appear more frequently in the dataset are represented with larger sizes in the visualization.

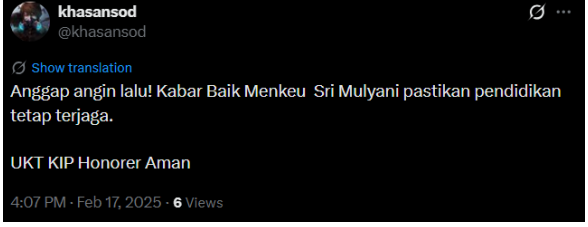
It is important to emphasize that word clouds are not used to determine sentiment directly. Instead, they serve as a complementary exploratory tool to provide contextual insights into the main issues and themes discussed by the public. Accordingly, the word cloud results are utilized to support the interpretation of sentiment classification outcomes produced by the IndoBERTweet model.

Word Cloud Sri Mulyani

Figure 23. Word Cloud Sri Mulyani

In the case of Sri Mulyani Indrawati, the word cloud results show clear differences across sentiment categories. For positive sentiment, dominant words such as “Sri Mulyani,” “Menkeu,” “aman,” “honorer,” “KIP,” and “pendidikan” reflect public perceptions that tend to support fiscal stability policies, education assistance programs, and trust in the performance of the Minister of Finance. This pattern is illustrated, for example, in a tweet posted by the account @khasansod on February 17, 2025.

Figure 24. Example of Positive Sentiment Sri Mulyani



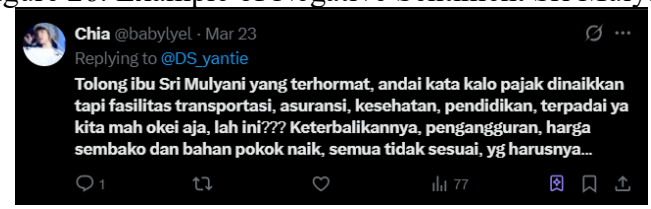
In the neutral sentiment category, the word cloud is dominated by informative terms such as “menteri keuangan,” “keuangan Sri,” “Indonesia,” and “pemerintah.” This indicates that a large portion of the conversation consists of news reporting or informational content without strong emotional polarity,

as seen in a tweet uploaded by the account @AgilZal on May 29, 2025.

Figure 25. Example of Neutral Sentiment Sri Mulyani

Meanwhile, in the negative sentiment category, words such as “pajak,” “utang,” “rakyat,” “naik,” and “ekonomi” appear most frequently. These terms reflect public criticism of fiscal policies, particularly those perceived to have direct economic impacts on society, as exemplified by a tweet posted on March 23, 2025 by the account @babylyel.

Figure 26. Example of Negative Sentiment Sri Mulyani

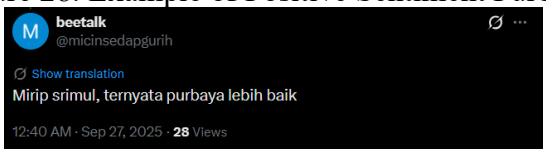


Word Cloud Purbaya

Figure 27. Word Cloud Purbaya

In contrast to Sri Mulyani Indrawati, the word cloud results for Purbaya Yudhi Sadewa exhibit a more limited and less diverse pattern. In the positive sentiment category, words such as “Purbaya,” “support,” “good,” and “excise” appear, although with relatively lower frequency. A tweet posted by the account @micinsedapgurih on September 27, 2025 reflects positive sentiment toward Purbaya.

Figure 28. Example of Positive Sentiment Purbaya



For neutral sentiment, the Purbaya word cloud is dominated by identity- and position-related terms such as “Purbaya Yudhi Sadewa,” “Board of Commissioners,” and “LPS.” The prominence of these terms indicates that most conversations are informational and descriptive in nature, as illustrated by a post from the account @Media_Asuransi on April 10, 2025.

Figure 29. Example of Positive Sentiment Purbaya

In the negative sentiment category, words such as “cigarettes,” “excise,” “omon,” “style,” and “economic growth” are dominant. These keywords reflect policy-specific criticism, particularly concerning issues related to economic growth and cigarette excise policies, as seen in a tweet posted by @PreciosaKanti on September 26, 2025.

Figure 30. Negative of Positive Sentiment Purbaya

Overall, the differences between the results for Sri Mulyani Indrawati and Purbaya Yudhi Sadewa indicate that data volume, vocabulary diversity, and the level of public exposure to each figure significantly influence sentiment analysis performance. The word cloud analysis strengthens the quantitative evaluation by providing a visual representation of dominant issues and narratives for each public figure. Consequently, the combination of metric-based evaluation and word cloud analysis offers a more comprehensive understanding of public sentiment dynamics as well as the limitations of the model when applied to imbalanced datasets.

CONCLUSION

Based on the results of this study, it can be concluded that the application of the IndoBERTweet model for sentiment analysis of public opinion on the X (Twitter) platform is able to produce reasonably good classification performance, particularly for public figures with sufficient data volume. This study examined two subjects, Sri Mulyani Indrawati and Purbaya Yudhi Sadewa, who differ significantly in terms of data characteristics and volume.

In the case of Sri Mulyani Indrawati, approximately 11,000 tweets were collected within the period from January 1, 2025 to November 1, 2025, with a maximum limit of 1,000 tweets per month. However, due to hardware limitations, the processed dataset was restricted to 2,800 tweets randomly selected using a Python script. These data were then used for the labeling, training, and testing stages of the model. With a relatively large and diverse dataset, the IndoBERTweet model demonstrated stable performance, achieving an accuracy of 82.45% and an F1-score of 82.69%. These results indicate that the model was able to capture positive, neutral, and negative sentiment contexts in a balanced manner, particularly within the neutral class, which was dominated by informational content.

In contrast, the dataset for Purbaya Yudhi Sadewa was much more limited, consisting of only approximately 650 tweets. This limitation directly affected model performance, which resulted in an

accuracy of 71.97% and an F1-score of 72.19%. Although these results can still be considered acceptable, the imbalance in data distribution across sentiment classes caused the model to be less optimal in consistently identifying sentiment patterns. This condition is further supported by the confusion matrix and classification report, which show lower precision and recall values for the negative sentiment class.

Beyond quantitative evaluation, word cloud analysis provided qualitative insights into the focus of public discourse. For Sri Mulyani Indrawati, public sentiment was largely associated with issues such as fiscal policy, taxation, education, and economic stability. Meanwhile, discussions related to Purbaya Yudhi Sadewa tended to focus on occupational identity and specific issues, particularly cigarette excise policy and economic growth. These findings suggest that public exposure and issue diversity significantly influence both sentiment distribution and sentiment analysis model performance.

Overall, this study concludes that IndoBERTweet is effective for Indonesian-language sentiment analysis; however, its performance is strongly influenced by data volume, class balance, and contextual diversity within the dataset. This research is not intended to be directly submitted to the government as a Final Project. Instead, the Final Project serves as a scientific foundation and proof of concept. Dissemination to governmental stakeholders may be pursued through academic publications, the development of public opinion monitoring systems, or further research conducted through institutional collaboration.

REFERENCES

[1] Juleha Juleha, Jusfira Yuniar, and Nur Riswandi Marsuki, "Peran Media Sosial Dalam Dinamika Opini Publik dan Partisipasi Politik Era Digital," *Concept: Journal of Social Humanities and Education*, vol. 3, no. 1, pp. 38–45, Jan. 2024, doi: 10.55606/concept.v3i1.951.

[2] F. Koto, J. H. Lau, and T. Baldwin, "IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Stroudsburg, PA, USA: Association for Computational Linguistics, Sep. 2021, pp. 10660–10668. doi: 10.18653/v1/2021.emnlp-main.833.

[3] U. Khairani, V. Mutiawani, and H. Ahmadian, "Pengaruh Tahapan Preprocessing Terhadap Model Indobert Dan Indobertweet Untuk Mendeteksi Emosi Pada Komentar Akun Berita Instagram," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 11, no. 4, pp. 887–894, Aug. 2024, doi: 10.25126/jtiik.1148315.

[4] U. Khairani, "PERBANDINGAN MODEL INDOBERT DAN INDOBERTWEET UNTUK MENDETEKSI EMOSI KOMENTAR BERITA PADA MEDIA SOSIAL TUGAS AKHIR," Jan. 2024. Accessed: Dec. 02, 2025. [Online]. Available: https://cdn.ar-raniry.ac.id/aps-test/teknologi_informasi/Pembimbing_di_USK_a.n_Hendri_Tugas_Akhir_Ulfia_Khairani_190810701006 8.pdf

[5] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 5, no. 2, p. 406, Apr. 2021, doi: 10.30865/mib.v5i2.2835.

[6] B. Eka Sarah Dewi, "SENTIMENT ANALYSIS MODEL ON ELECTRIC VEHICLES USING INDOBERTWEET AND INDOBERT ALGORITHM," *ANTIVIRUS: Jurnal Ilmiah Teknik Informatika*, 2025, doi: https://doi.org/10.35457/w5r3g517.

[7] J. C. Setiawan, K. M. Lhaksmana, and B. Bunyamin, "Sentiment Analysis of Indonesian TikTok Review Using LSTM and IndoBERTweet Algorithm," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, pp. 774–780, Aug. 2023, doi: 10.29100/jipi.v8i3.3911.

[8] K. Kadarsih and D. Pujianto, "APPLICATION OF TRANSFORMER MODEL AND WORD EMBEDDING IN SENTIMENT ANALISYS OF INDONESIAN E-COMMERCE APPLICATION REVIEW," *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 7, no. 3, pp. 720–732, Jul. 2025, doi: 10.47709/cnahpc.v7i3.6354.

[9] N. M. Damayanti, "ANALISIS SENTIMEN PUBLIK PADA TAGAR #BTSCOMEBACK DI PLATFORM X MENGGUNAKAN INDOBERTWEET," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 3, Jul. 2025, doi: 10.23960/jitet.v13i3.7176.

[10] J. F. Kusuma and A. Chowanda, "Indonesian Hate Speech Detection Using IndoBERTweet and BiLSTM on Twitter," *JOIV : International Journal on Informatics Visualization*, vol. 7, no. 3, pp. 773–780, Sep. 2023, doi: 10.30630/joiv.7.3.1035.

[11] Z. Alhaq, A. Mustopa, S. Mulyatun, and J. D. Santoso, "PENERAPAN METODE SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN PENGGUNA TWITTER," *Journal of Information System Management (JOISM)*, vol. 3, no. 2, pp. 44–49, Jul. 2021, doi: 10.24076/joism.2021v3i2.558.