Fine-tuning SahabatAI for Sentiment Analysis of Indonesian Ministry of Foreign Affairs Communications and Public Discourse

Abstract

This thesis proposal outlines a research plan to investigate the efficacy of fine-tuning SahabatAI, a novel Large Language Model (LLM) specifically pre-trained on Bahasa Indonesia, for sentiment analysis within the domain of Indonesian foreign policy. The study will focus on communications from the Ministry of Foreign Affairs of the Republic of Indonesia (Kemlu RI) and related public discourse, particularly concerning Indonesia's ASEAN Chairmanship in 2023. The research will involve a comprehensive review of state-of-the-art sentiment analysis techniques, with an emphasis on Transformer-based models like IndoBERT and multilingual counterparts, to establish a comparative baseline. A detailed methodology for data collection from official Kemlu RI sources, Indonesian news media (Antara, Kompas, Tempo, Detik), and potentially Twitter/X will be presented, followed by a robust plan for data preprocessing, annotation, and the fine-tuning of a SahabatAI base model using Parameter-Efficient Fine-Tuning (PEFT) techniques like LoRA/QLoRA. Performance will be evaluated using standard classification metrics and qualitative analysis. This research aims to contribute to the advancement of Indonesian NLP resources, provide insights into SahabatAI's capabilities for specialized governmental applications, and offer a potential tool for Kemlu RI to gauge public and media sentiment regarding its diplomatic initiatives. The project is designed for completion within a six-month timeframe.

Chapter 1: Introduction

1.1. Background

The proliferation of digital information has profoundly impacted international relations and diplomatic practice. Ministries of Foreign Affairs (MFAs) worldwide are increasingly confronted with the need to monitor and analyze vast quantities of online text, including official communications, media coverage, and public opinion expressed on social media, to effectively manage their country's image and formulate responsive foreign policies.[1] Sentiment analysis, a subfield of Natural Language Processing (NLP), offers a powerful means to automatically extract subjective information—opinions, emotions, and attitudes—from textual data, providing valuable insights for diplomatic strategy and public diplomacy efforts.[3]

For Indonesia, a key player in ASEAN and global affairs, understanding domestic and international sentiment towards its foreign policy is crucial. The Ministry of Foreign Affairs of the Republic of Indonesia (Kementerian Luar Negeri Republik Indonesia, hereafter Kemlu RI) actively engages in public diplomacy and utilizes digital platforms for communication.[6] Analyzing the sentiment expressed in relation to Kemlu RI's initiatives, such as its chairmanship of ASEAN in 2023, can provide actionable intelligence for policy refinement and communication strategies.[8]

Recent advancements in NLP, particularly the development of Large Language Models (LLMs) and Transformer-based architectures like BERT [9], RoBERTa [9], and XLM-R [9], have significantly improved the performance of sentiment analysis tasks. For Bahasa Indonesia, models like IndoBERT [10] have demonstrated strong capabilities. More recently, SahabatAI, an LLM specifically pre-trained with a significant corpus of Bahasa Indonesia and its regional dialects, has emerged, presenting a new opportunity to develop highly effective, locally-tuned NLP solutions.[11]

1.2. Problem Statement

While general-purpose sentiment analysis tools and multilingual models exist, their performance can be suboptimal when applied to specialized domains like diplomatic language or low-resource languages such as Bahasa Indonesia without specific fine-tuning.[16] Diplomatic texts often contain nuanced language, subtle expressions of stance, and context-dependent sentiment that generic models may fail to capture accurately.[3] Furthermore, the unique linguistic characteristics of Bahasa Indonesia, including colloquialisms, sarcasm (a known challenge in sentiment analysis [18]), and evolving terminology, necessitate models specifically adapted to this context.

The advent of SahabatAI [11] offers a promising avenue to address these challenges. However, its efficacy for domain-specific sentiment analysis, particularly within the Indonesian governmental and foreign policy context, has yet to be extensively benchmarked against established Indonesian NLP models like IndoBERT or robust multilingual models. There is a need to explore how effectively SahabatAI can be fine-tuned for this specific task and whether it offers performance advantages due to its extensive Indonesian pre-training.

1.3. Research Questions (RQs)

This research aims to answer the following questions:

- RQ1: How effectively can a SahabatAI base model be fine-tuned for sentiment analysis (classifying text as Positive, Negative, or Neutral) on a dataset comprising Kemlu RI communications and related public/media discourse

concerning Indonesia's 2023 ASEAN Chairmanship?
- RQ2: What is the performance of the fine-tuned SahabatAI model compared to a fine-tuned IndoBERT model and potentially other relevant Transformer-based models when evaluated on the same domain-specific Indonesian sentiment analysis task?
- RQ3: What are the optimal data collection, preprocessing, and annotation strategies for creating a suitable Indonesian sentiment analysis dataset for the foreign policy domain within the constraints of a Master's thesis?
- RQ4: What are the key linguistic challenges (e.g., nuance, sarcasm, officialese) encountered when performing sentiment analysis on Indonesian diplomatic and media texts, and how well does the fine-tuned SahabatAI model address them?

1.4. Research Objectives (ROs)

The objectives of this research are:

- RO1: To develop a methodology for collecting and preparing a domain-specific dataset in Bahasa Indonesia for sentiment analysis, focusing on Kemlu RI communications and public/media discourse related to Indonesia's 2023 ASEAN Chairmanship.
- RO2: To fine-tune a selected SahabatAI base model (Llama3 or Gemma2 variant) for sentiment classification using the prepared dataset and Parameter-Efficient Fine-Tuning (PEFT) techniques.
- RO3: To evaluate the performance of the fine-tuned SahabatAI model using standard classification metrics (Accuracy, Precision, Recall, F1-score) and compare it against a fine-tuned IndoBERT model and potentially other relevant baselines.
- RO4: To conduct a qualitative error analysis to identify strengths and weaknesses of the fine-tuned SahabatAI model in interpreting sentiment in Indonesian diplomatic and media texts.
- RO5: To document the process and provide insights into the feasibility and effectiveness of using SahabatAI for specialized governmental NLP tasks in Indonesia.

1.5. Significance of the Study

This research holds significance in several areas:

- **Advancement of Indonesian NLP:** It will contribute to the growing body of research on NLP for Bahasa Indonesia by evaluating a new, large-scale Indonesian LLM (SahabatAI) on a specialized task. The creation of a domain-specific annotated dataset (even if modest in size) can be a valuable resource.

- **Practical Application for Kemlu RI:** The findings can inform Kemlu RI about the potential of using advanced AI tools like SahabatAI for automated sentiment analysis, aiding in public diplomacy, media monitoring, and understanding public perception of foreign policy initiatives.
- **Benchmarking SahabatAI:** This study will provide one ofthe early benchmarks for SahabatAI's fine-tuning performance on a specific downstream task, offering insights for other researchers and practitioners looking to leverage this model.
- **Technical Contribution:** The research will explore practical aspects of fine-tuning large Indonesian LLMs with limited resources, focusing on PEFT techniques, which is relevant for academic and applied research in similar contexts.

1.6. Scope and Limitations

- **Scope:**
  - The primary focus is on sentiment analysis (Positive, Negative, Neutral classification) of Indonesian texts.
  - The domain is Kemlu RI communications and related public/media discourse, specifically concerning Indonesia's ASEAN Chairmanship in 2023, to ensure a manageable scope for data collection and analysis within a 6-month thesis.
  - The primary model for investigation is SahabatAI (a base Llama3 or Gemma2 variant). IndoBERT will serve as a key baseline.
  - Data sources will include official Kemlu RI publications, selected Indonesian news media, and potentially Twitter/X.
- **Limitations:**
  - **Dataset Size:** Due to the time constraints of a Master's thesis, the manually annotated dataset for fine-tuning will likely be of limited size. This may impact the generalizability of the fine-tuned model.
  - **Computational Resources:** Access to high-performance computing resources will be necessary but may be limited, influencing the scale of experimentation (e.g., extent of hyperparameter tuning, number of models compared). PEFT methods are chosen to mitigate this.
  - **API Access:** Data collection from news APIs (e.g., NewsAPI.org) or Twitter/X API is subject to availability, cost, and policy changes, which could affect data acquisition.[20]
  - **Subjectivity of Sentiment:** Sentiment annotation can be subjective, especially for nuanced diplomatic language. Efforts will be made to ensure consistency through clear guidelines, but inherent subjectivity remains a limitation.

    - **Focus on Text:** This study will focus on textual sentiment analysis and will not incorporate multimodal data (images, videos).

1.7. Thesis Structure

This thesis will be organized as follows:

- **Chapter 1: Introduction:** Provides the background, problem statement, research questions, objectives, significance, scope, and limitations.
- **Chapter 2: Literature Review and Related Work:** Discusses SOTA in sentiment analysis (general and Indonesian), Transformer models, AI in diplomacy, challenges in Indonesian NLP, and relevant prior studies.
- **Chapter 3: Methodology:** Details the chosen research topic (sentiment analysis of MoFA communications using SahabatAI), explains the SahabatAI model, justifies its selection, and outlines the plan for data collection, preparation (including annotation), model fine-tuning (SahabatAI and baselines), and evaluation metrics.
- **Chapter 4: Implementation Plan and Timeline:** Presents a detailed 6-month work plan, lists required resources, and conducts a risk assessment with mitigation strategies.
- **Chapter 5: Expected Outcomes and Dissemination:** Discusses anticipated contributions, potential future research directions, and plans for disseminating the research findings.
- **References:** Lists all cited sources.
- **Appendices (if any):** May include detailed annotation guidelines or sample data.

Chapter 2: Literature Review and Related Work

2.1. Sentiment Analysis: Techniques and Challenges

Sentiment analysis, or opinion mining, is the computational study of opinions, sentiments, and emotions expressed in text.[16] It aims to classify the polarity of a given text as positive, negative, or neutral, and can extend to more granular emotions (e.g., joy, anger, fear).[25] Methodologies broadly fall into lexicon-based approaches, traditional machine learning (ML) approaches, and deep learning (DL) approaches.[27]

- **Lexicon-based methods** rely on sentiment dictionaries (lexicons) where words are assigned polarity scores (e.g., SentiWordNet, InSet for Indonesian [29]). While simple to implement, they often struggle with context, negation, and domain-specific sentiment expressions.[27] Studies on Indonesian sentiment using lexicons like InSet have shown moderate accuracy, often serving as a baseline or a component in hybrid systems.[28]

- **Traditional ML methods** like Naïve Bayes (NB) and Support Vector Machines (SVM) use features extracted from text (e.g., TF-IDF, n-grams) to train classifiers.[33] SVMs have shown good performance in Indonesian sentiment tasks.[33] For instance, a study on Indonesian election-related tweets using SVM reported high accuracy [35], and another comparing SVM with SentiWordNet for Indonesian text found SVM superior, especially with balanced datasets.[34] NB has also been applied, for example, to analyze sentiment towards the Indonesian Ministry of Investment, achieving 82% accuracy with a rule-based approach.[24]
- **Deep Learning (DL) methods**, particularly those based on Recurrent Neural Networks (RNNs, LSTMs, GRUs) and Transformer architectures (e.g., BERT), have become state-of-the-art.[33] These models can learn complex patterns and contextual representations from large amounts of data.

Challenges in Sentiment Analysis:
Common challenges include handling negation, sarcasm, irony, ambiguity, and domain-specific language.5 For diplomatic language, these challenges are amplified due to its inherent nuance, indirectness, and strategic use of ambiguity.3 Data scarcity, especially for low-resource languages like Bahasa Indonesia and specific domains like foreign policy, is another significant hurdle.16 Colloquialisms and informal language prevalent in social media also pose difficulties for Indonesian NLP.18

2.2. Transformer-Based Models for Sentiment Analysis

Transformer models [41], introduced by Vaswani et al. (2017), have revolutionized NLP. Their self-attention mechanism allows them to weigh the importance of different words in a sequence, capturing long-range dependencies and contextual information effectively.

- **BERT (Bidirectional Encoder Representations from Transformers)** [9] and its variants (e.g., RoBERTa [9]) are pre-trained on massive text corpora and can be fine-tuned for various downstream tasks, including sentiment analysis, achieving SOTA results.[36]
- **Multilingual Models:** mBERT (multilingual BERT) [41] and XLM-RoBERTa (XLM-R) [41] are pre-trained on text from many languages and have shown strong cross-lingual transfer capabilities. They can be effective for low-resource languages where monolingual pre-trained models are scarce.[41] However, their performance on a specific language might be surpassed by a well-trained monolingual model for that language.
- **Indonesian-Specific Models:**
    - **IndoBERT** [42]: Developed by Koto et al. (2020) and Wilie et al. (2020), IndoBERT models are pre-trained on large Indonesian corpora (Indo4B). They have demonstrated superior performance on various Indonesian NLP tasks,

including sentiment analysis within the IndoNLU benchmark, often outperforming multilingual models.[10] The IndoNLU benchmark includes the SmSA (Sentence-level Sentiment Analysis) dataset, a collection of comments and reviews from multiple online platforms, labeled as positive, negative, or neutral.[10] Fine-tuned IndoBERT models, such as taufiqdp/indonesian-sentiment [30] and crypter70/IndoBERT-Sentiment-Analysis [44], have achieved high accuracy on general Indonesian sentiment tasks.
  - **SahabatAI** [11]: A recent initiative by GoTo Group and Indosat Ooredoo Hutchison, in collaboration with AI Singapore and others.[11] SahabatAI provides a collection of LLMs (based on Llama3 and Gemma2 architectures) that have undergone continued pre-training on approximately 50 billion tokens, with a significant portion being Indonesian text (including SEA-LION Pile and other Indonesian sources) and also including Javanese and Sundanese dialects.[11] Both base and instruct-tuned versions are available.[11] Benchmark results on the SEA HELM (BHASA) evaluation suite show strong performance on sentiment analysis tasks for Indonesian, Javanese, and Sundanese, often outperforming other contemporary models in zero-shot instruct settings.[11] The availability of these powerful, Indonesian-centric base models offers a new opportunity for fine-tuning on specific downstream tasks like domain-specific sentiment analysis.

2.3. AI in Diplomacy and Public Opinion Analysis

AI is increasingly being applied in diplomacy for various purposes:

- **Decision Support & Trend Analysis:** AI tools analyze large datasets to identify geopolitical trends and predict developments, aiding diplomats in formulating informed policies.[1]
- **Drafting Documents & Communication:** AI can assist in drafting speeches and official documents.[49]
- **Negotiation Support:** AI can analyze past negotiations and identify areas of potential agreement or conflict.[1]
- **Crisis Management & Early Warning:** Predictive analytics and real-time information gathering help in crisis response and conflict forecasting.[1] Projects like the Bilateral Navigator use AI to analyze relationships between UN member states.[50]
- **Public Diplomacy & Media Monitoring:** Sentiment analysis of media and public opinion helps MFAs understand how their country and policies are perceived, track narratives, and detect misinformation.[49] A study on China's MFA statements

regarding the Russo-Ukrainian war used NLP for topic modeling and sentiment analysis (pro-Russia, neutral, pro-Ukraine).[54] Another analyzed UN speeches on the same conflict using VADER and BERT.[3]

Kemlu RI itself has been leveraging ICT for public communication and citizen services (e.g., SafeTravel app, Peduli WNI portal) [6], and there's an acknowledged need to refine its digital diplomacy strategy.[7] Analyzing media framing of Indonesian foreign policy is also an area of academic interest.[55]

2.4. Related Work in Indonesian Sentiment Analysis (Political/Diplomatic Context)

While general Indonesian sentiment analysis research is growing [57], studies specifically focusing on the political or foreign policy domain are less common but emerging.

- A study analyzed sentiment towards the Indonesian Ministry of Investment using Naïve Bayes.[24]
- Research on the 2024 Indonesian presidential election used YouTube comments for multi-label emotion classification, employing GPT-3.5 for initial labeling and Bi-LSTM for classification.[25]
- Another study on the Kanjuruhan tragedy used SVM to analyze sentiment from tweets and YouTube comments, with Indonesian RoBERTa used for labeling.[35]
- Ferdiana et al. (2019) presented a general Indonesian sentiment dataset from Twitter and benchmarked SVM, KNN, and SGD.[61]
- Research on Indonesian public diplomacy often discusses information dissemination and image management but may not always employ computational sentiment analysis.[62]
- A study by Jurnal Dinamika Global used framing analysis (a qualitative method) to assess Indonesian mass media coverage of foreign policy concerning US-China rivalry, concluding media support for Indonesia's neutral stance.[55] This highlights the relevance of media analysis for foreign policy but also the opportunity for quantitative sentiment analysis.
- A study analyzed international media and Twitter perceptions of Indonesia, using aspect-based sentiment analysis for news and classification algorithms for Twitter, noting that diplomacy and export aspects generally received positive sentiment.[8] This study also highlighted limitations due to insufficient training data for Indonesian content.

The current proposal aims to build upon this by applying a new, powerful Indonesian LLM (SahabatAI) to the specific domain of Kemlu RI communications and related media discourse, focusing on a significant diplomatic event (ASEAN Chairmanship

2023), and providing a direct comparison with established models like IndoBERT. The emphasis on fine-tuning a base SahabatAI model for this classification task, rather than relying solely on instruct models or zero-shot prompting, is a key technical focus.

Chapter 3: Methodology - Fine-tuning SahabatAI for Sentiment Analysis of MoFA Communications

This chapter details the methodology for fine-tuning a SahabatAI base model to perform sentiment analysis on texts related to the Indonesian Ministry of Foreign Affairs (Kemlu RI) and its 2023 ASEAN Chairmanship.

3.1. Chosen Technique: Fine-tuning SahabatAI for Sentiment Classification

The core technique will be **supervised fine-tuning of a SahabatAI base model for text classification**. Sentiment analysis will be framed as a multi-class classification task, predicting one of three labels: **Positive, Negative, or Neutral**.

SahabatAI Overview:
SahabatAI is a collection of LLMs co-initiated by GoTo Group and Indosat Ooredoo Hutchison, with development by GoTo and AI Singapore.11 It includes models based on Llama3 and Gemma2 architectures, which have undergone continued pre-training on approximately 50 billion tokens. A significant portion of this pre-training data is in Bahasa Indonesia (including the SEA-LION Pile, Indonesian Wikipedia, and news sources), with additional data in Javanese, Sundanese, and English.11 Both base and instruction-tuned versions are available on Hugging Face.11 For this research, a base version of SahabatAI (either GoToCompany/llama3-8b-cpt-sahabatai-v1-base 11 or GoToCompany/gemma2-9b-cpt-sahabatai-v1-base 12) will be selected.

Justification for Selecting SahabatAI (Base Model):

1. **Indonesian-Centric Pre-training:** SahabatAI's extensive pre-training on a large and diverse Indonesian corpus is its primary advantage. This should provide a strong foundation for understanding the nuances, vocabulary, and grammatical structures of Bahasa Indonesia, potentially leading to better performance on downstream tasks compared to multilingual models or models with less Indonesian pre-training data.[11]
2. **State-of-the-Art Architecture:** Being based on Llama3 and Gemma2 architectures, SahabatAI incorporates recent advancements in LLM design, offering potentially superior capabilities compared to older models like the original BERT architecture used in early IndoBERT versions.[11]
3. **Suitability of Base Models for Classification Fine-tuning:** Base LLMs, which are pre-trained to predict the next token, are generally more flexible for fine-tuning on specific downstream tasks like classification when a custom

labeled dataset is available. Fine-tuning a base model typically involves adding a classification head on top of the pre-trained model and training these components on the labeled data. This is a standard and effective approach for adapting LLMs to classification tasks.[66] Instruct-tuned models, while powerful for instruction following, might require framing the sentiment task as an instruction, which may not be the most direct or optimal method for this specific classification problem and adds complexity.[66]

4. **Opportunity for Novel Benchmarking:** As SahabatAI is relatively new, this research provides an opportunity to benchmark its fine-tuning performance on a specialized, real-world governmental task, contributing valuable insights to the Indonesian NLP community.

Comparison with Other Approaches:

- **IndoBERT:** IndoBERT (e.g., indobenchmark/indobert-base-p1) is a strong baseline as it was specifically pre-trained for Indonesian and has shown excellent performance on the IndoNLU benchmark, including sentiment analysis (SmSA task).[42] Fine-tuning IndoBERT on the same custom dataset will provide a direct comparison to assess SahabatAI's relative performance.
- **Multilingual Models (e.g., XLM-R):** While models like XLM-R perform well cross-lingually and can be effective for low-resource languages [41], a dedicated monolingual model with extensive in-language pre-training like SahabatAI or IndoBERT is often expected to perform better on monolingual tasks for that specific language.
- **Lexicon-based and Traditional ML:** These methods generally underperform compared to Transformer-based models on complex sentiment tasks, especially with nuanced language, but an SVM with TF-IDF might serve as an additional, simpler baseline if time permits.[33]

The primary hypothesis is that SahabatAI, due to its recent architecture and extensive Indonesian pre-training, will achieve competitive or superior performance compared to IndoBERT when fine-tuned on the domain-specific sentiment analysis task.

3.2. Data Collection Plan

The dataset will focus on texts related to Kemlu RI and Indonesia's ASEAN Chairmanship in 2023. This specific focus provides a defined scope and allows for the collection of domain-relevant data.

Data Sources:

1. **Kemlu RI Official Communications:**

- **Source:** Official Kemlu RI website (kemlu.go.id).
- **Content:** Press releases (Siaran Pers) [31], speeches (Pernyataan Pers Tahunan Menteri Luar Negeri - PPTM, other official statements) [70], and relevant reports or news articles published by Kemlu RI. The Ministry's regulations indicate that official websites and social media are used to disseminate information on foreign policy.[73]
- **Collection Method:** Manual download of PDFs and web scraping of HTML content using Python libraries (e.g., requests, BeautifulSoup4). PDF text extraction will be performed using libraries like PyPDF2 or pdfminer.six.
- **Keywords for Filtering (within Kemlu sources):** "ASEAN Chairmanship 2023", "Keketuaan ASEAN RI 2023", "KTT ASEAN 2023" (ASEAN Summit 2023), names of key policy initiatives or outcomes related to the chairmanship.
- **Time Period:** Primarily documents published from late 2022 (covering preparations) through early 2024 (covering outcomes and follow-ups related to the 2023 chairmanship).

2. **Indonesian News Media:**
   - **Selected Outlets:** Major Indonesian news providers known for national and international coverage. Based on previous research and availability, these will include:
     - Antara News (Kantor Berita Antara - LKBN Antara) [74]
     - Kompas.com / Kompas.id [55]
     - Tempo.co [55]
     - Detik.com [81]
   - **Collection Method:**
     - **NewsAPI.org:** Explore using NewsAPI.org (Developer plan if feasible, or free tier for limited tests) to gather articles by specifying Indonesian sources and keywords.[75] Source IDs for Indonesian outlets would need to be confirmed via the /sources endpoint or their documentation.[75]
     - **Kompas.id API:** Kompas.id offers an API for partners, which might be explored if direct access is possible, though this is unlikely for a short-term academic project.[77]
     - **Web Scraping (Primary method for news):** If API access is limited or costly, targeted web scraping using Python (requests, BeautifulSoup4, potentially Scrapy) will be the primary method. This will focus on the archive/search sections of the news websites.
   - **Keywords for Filtering (News):** "Keketuaan ASEAN Indonesia 2023", "Kemlu RI ASEAN", "Menlu Retno Marsudi ASEAN 2023", specific ASEAN summit names held in Indonesia in 2023, key policy themes discussed during the chairmanship (e.g., "Indo-Pasifik ASEAN", "ekonomi digital ASEAN").

- **Time Period:** News articles published during 2023, with some coverage from late 2022 and early 2024.

3. **Twitter/X (Optional, Contingent on API Access and Time):**
   - **Source:** Publicly available tweets in Bahasa Indonesia.
   - **Content:** Reactions and opinions regarding Kemlu RI's handling of the ASEAN Chairmanship.
   - **Collection Method:** Twitter API v2. Given the current pricing and access limitations [20], access will likely be limited to the **Basic tier** ($200/month, 15,000 posts/month) [22] or the very restrictive **Free tier** (100 posts/month).[22] Academic Research access track changes mean it's no longer a high-volume free option.[23] Libraries like Tweepy [89] would be used if API access is secured.
   - **Keywords/Filters (Twitter):** Hashtags like #ASEANIndonesia2023, #KeketuaanASEANRI, relevant keywords from MoFA statements regarding the chairmanship, and mentions of @Kemlu_RI or key officials in the context of ASEAN 2023.
   - **Filtering:** Language will be filtered to Bahasa Indonesia (lang:id). Geolocation filtering (Indonesia) will be applied if the API allows and it is deemed relevant.
   - **Time Period:** Tweets posted during key periods of the ASEAN Chairmanship in 2023 (e.g., around major summits or announcements).
   - **Anticipated Volume:** Highly dependent on the API access tier secured. If limited to the Basic tier, collection will be highly focused on specific event windows to maximize relevant data within the 15,000 post limit.

The collection of Twitter data is contingent on securing adequate API access. Given the current restrictions and costs associated with the Twitter API [20], this data source might be limited in scope or, in a worst-case scenario, excluded if feasible access cannot be obtained. This represents a significant constraint that must be addressed early in the research.

Table 3.1: Overview of Data Sources for MoFA Sentiment Analysis (ASEAN Chairmanship 2023 Focus)

Data Source	Content Type	Collection Method	Estimated Volume (Focus on ASEAN Chairmanship 2023)	Time Period	Keywords/Filters (Examples)
Kemlu RI	Press	Web	200-500	Late 2022 -	"ASEAN

Website	Releases, Speeches, Official Reports	Scraping (Python: BeautifulSoup, Scrapy), PDF Text Extraction	documents	Early 2024	Chairmanship 2023", "Keketuaan ASEAN RI", specific summit names, policy themes (e.g., "Indo-Pacific ASEAN")
Antara News	News Articles	NewsAPI.org (if source ID available & feasible tier), Web Scraping	200-600 articles	2023 (late 2022-early 2024)	"Keketuaan ASEAN Indonesia 2023", "Kemlu ASEAN", names of key officials + ASEAN
Kompas.com	News Articles	Kompas.id API (explore feasibility), NewsAPI.org (if available), Web Scraping	200-600 articles	2023 (late 2022-early 2024)	"Keketuaan ASEAN Indonesia 2023", "Kemlu ASEAN", names of key officials + ASEAN
Tempo.co	News Articles	NewsAPI.org (if available), Web Scraping	200-600 articles	2023 (late 2022-early 2024)	"Keketuaan ASEAN Indonesia 2023", "Kemlu ASEAN", names of key officials + ASEAN
Detik.com	News Articles	NewsAPI.org (if available), Web	200-600 articles	2023 (late 2022-early 2024)	"Keketuaan ASEAN Indonesia

		Scraping			2023", "Kemlu ASEAN", names of key officials + ASEAN
Twitter/X (Kemlu)	Tweets (Public Reactions)	Twitter API v2 (Basic tier if feasible)	Max 15,000 (if Basic tier) or less depending on budget/focus	Key periods in 2023	#ASEANIndonesia2023, @Kemlu_RI + ASEAN, relevant policy keywords; lang:id

3.3. Dataset Construction and Annotation

3.3.1. Data Cleaning and Preprocessing

Once the raw textual data is collected, a rigorous cleaning and preprocessing pipeline will be implemented to prepare it for sentiment analysis. This pipeline will include:

- **Standard Text Preprocessing:**
    - Conversion to lowercase to ensure consistency.
    - Removal of HTML tags, URLs, email addresses, and other web-specific artifacts using regular expressions or specialized libraries.
    - Removal of special characters, excessive punctuation (while retaining sentence-ending punctuation and potentially sentiment-indicative punctuation like exclamation marks), and digits not relevant to sentiment. These are common steps in preparing text for NLP tasks.[5]
- **Indonesian-Specific Preprocessing:**
    - **Normalization of Slang and Colloquial Words:** For text from news comments or social media, informal language, slang (bahasa alay), and abbreviations are common. A dictionary of Indonesian slang words (e.g., leveraging resources like Kamus Alay from sources like [58]) will be used for normalization. Custom mappings may be developed based on observed frequent informal terms in the collected data. This step is crucial for standardizing vocabulary.
    - **Stemming:** The Sastrawi library [57], a popular stemmer for Bahasa Indonesia, will be used to reduce words to their root form (e.g., "kebijakan" -> "bijak").

This helps in consolidating related word forms.
  - **Stopword Removal:** A comprehensive list of Indonesian stopwords (common words like "yang", "di", "dan", "ini") will be used to filter out words that generally do not carry significant sentiment. This list will be sourced from standard Indonesian NLP resources [57] and potentially augmented with domain-specific stopwords identified from the MoFA corpus (e.g., frequently occurring but sentiment-neutral diplomatic boilerplate phrases).
  - **Tokenization:** Text will be tokenized into individual words or sub-word units as required by the SahabatAI tokenizer (which uses the default Llama3 or Gemma2 tokenizer).[11]

3.3.2. Sentiment Labeling Strategy

High-quality sentiment labels are paramount for effective fine-tuning. The primary sentiment categories will be: **Positive, Negative, and Neutral.**

- **Consideration for Aspect-Based Sentiment:** While document-level sentiment is the primary goal, if the collected data and research questions allow, an exploratory analysis of aspect-based sentiment could be considered. For example, sentiment towards specific *themes* of Indonesia's ASEAN Chairmanship (e.g., economic cooperation, regional security, Myanmar issue). This would significantly increase annotation complexity and may be deferred to future work if not feasible within the Master's thesis timeline.[8]
- Annotation Approach:
  Given the resource constraints of a Master's thesis and the time-intensive nature of manual annotation, a hybrid approach leveraging both automated techniques and human oversight is proposed. The finding that as few as 1000 high-quality observations can yield good results for transfer learning in Indonesian emotion classification is particularly relevant here.26
  1. **Initial Dataset Size:** Aim to create a fine-tuning dataset of approximately 2,000-5,000 instances, balanced across sources (MoFA documents, news articles, and a smaller, focused set of tweets if available).
  2. **Annotation Guideline Development:** Clear, detailed annotation guidelines will be developed. These guidelines will define each sentiment category with specific criteria and provide illustrative examples from the collected domain-specific texts (MoFA statements, news articles about foreign policy). This is crucial for consistency.[25]
  3. **Pilot Annotation & Guideline Refinement:** A small subset of data (e.g., 100-200 instances) will be independently annotated by at least two annotators (native Indonesian speakers, ideally with some understanding of

political or media discourse; the researcher will be one). Inter-annotator agreement (IAA) will be calculated (e.g., Cohen's Kappa). Disagreements will be discussed, and guidelines will be refined to improve clarity and consistency.

4. **Primary Annotation Strategy - Manual Annotation of a Core Set:** A core dataset of 1,000-2,000 instances will be manually annotated by the primary researcher, adhering strictly to the refined guidelines. This size is chosen based on the potential effectiveness of smaller, high-quality datasets for fine-tuning LLMs.[26]
5. **Augmentation/Expansion (if needed and feasible):**
   - **Semi-Supervised Labeling:** If a larger dataset is desired and manual annotation of the full set is too time-consuming, an existing high-performing Indonesian sentiment model (e.g., a fine-tuned IndoBERT like taufiqdp/indonesian-sentiment [30] or crypter70/IndoBERT-Sentiment-Analysis [44]) could be used to provide initial labels for an additional set of documents. These machine-generated labels would then be manually reviewed and corrected for a significant random sample to ensure quality.
   - **Lexicon-Based Pre-filtering (Exploratory):** Lexicons like InSet [29] could be used to pre-filter a larger pool of uncategorized text to identify potentially sentiment-rich documents for further manual annotation, though lexicon-based methods alone are not reliable enough for final labeling in this nuanced domain. The emphasis will be on the quality of annotations over sheer quantity, especially given the promising results from fine-tuning with smaller datasets in related Indonesian NLP tasks.[26]

Table 3.2: Sentiment Annotation Scheme and Examples (Illustrative)

Sentiment Label	Definition	Example Text (Bahasa Indonesia - Hypothetical, related to ASEAN Chairmanship)	Justification for Label
Positive	Expresses approval, support, optimism, satisfaction, achievement, or positive outcomes related to Kemlu RI's actions/statements or	"Keketuaan ASEAN Indonesia berhasil mendorong kesepakatan penting terkait ekonomi digital di kawasan, ini langkah maju yang	Highlights success and positive outcome ("berhasil mendorong", "langkah maju signifikan").

	Indonesia's role in the ASEAN Chairmanship.	signifikan." (Indonesia's ASEAN Chairmanship successfully pushed for an important agreement regarding the digital economy in the region, this is a significant step forward.)	
Negative	Expresses disapproval, criticism, pessimism, dissatisfaction, failure, or negative outcomes related to Kemlu RI's actions/statements or Indonesia's role in the ASEAN Chairmanship.	"Meskipun Indonesia memimpin ASEAN, penyelesaian isu Myanmar masih jalan di tempat dan belum menunjukkan hasil konkret." (Although Indonesia leads ASEAN, the resolution of the Myanmar issue is still stagnant and has not shown concrete results.)	Expresses criticism and lack of progress ("jalan di tempat", "belum menunjukkan hasil konkret").
Neutral	Provides factual information, reports on events without explicit positive or negative judgment, or expresses an ambivalent stance. Official statements detailing procedures or agendas often fall here.	"Pertemuan Tingkat Menteri Luar Negeri ASEAN akan diselenggarakan di Jakarta pada tanggal 10-12 Juli 2023 untuk membahas agenda prioritas keketuaan." (The ASEAN Foreign Ministers' Meeting will be held in Jakarta on July 10-12, 2023, to discuss the chairmanship's priority agenda.)	Factual reporting of an event and agenda without explicit sentiment.

3.3.3. Data Splitting

The final annotated dataset will be divided into three mutually exclusive sets:

- **Training Set (70%):** Used to fine-tune the SahabatAI model.
- **Validation Set (15%):** Used for hyperparameter tuning and monitoring the model's performance during training to prevent overfitting.
- **Test Set (15%):** Used for the final evaluation of the fine-tuned model's performance on unseen data.

The splits will be stratified to ensure that the proportion of each sentiment class is maintained across all three sets, which is important especially if class imbalances exist. An 80:10:10 split is also a common alternative.[32]

The annotation phase is recognized as a potential bottleneck. The strategy aims for a high-quality, albeit potentially smaller, manually annotated core dataset, drawing on findings that suggest effectiveness even with limited data for transfer learning.[26] This pragmatic approach balances the need for quality with the time constraints of a Master's project.

3.4. Model Fine-Tuning Strategy for SahabatAI

3.4.1. Selection of SahabatAI Base Model

The primary model for fine-tuning will be a **base version** of SahabatAI, rather than an instruct-tuned version. The choice will be between:

- GoToCompany/llama3-8b-cpt-sahabatai-v1-base [11]
- GoToCompany/gemma2-9b-cpt-sahabatai-v1-base [12]
- **Justification for Base Model Selection:**
  - **Task Suitability:** Sentiment analysis is fundamentally a classification task. Base LLMs are generally more adaptable for fine-tuning on specific downstream tasks like classification when a custom labeled dataset is available. They provide a more "neutral" foundation, allowing the model to learn the specific patterns and nuances of the sentiment task from the provided data without being overly influenced by prior instruction-following behaviors.[66]
  - **Flexibility:** Fine-tuning a base model for sequence classification typically involves adding a classification head on top of the pre-trained encoder/decoder stack and training these components on the labeled data. This is a standard and well-understood approach.[91]
  - **Avoiding Instruction Formatting Overhead:** Using an instruct model would necessitate framing the sentiment analysis task as an instruction (e.g., "Classify the sentiment of the following text: [text]. Options: Positive, Negative, Neutral."). While possible, this might not be the most natural or

effective way to leverage the model for this specific classification problem and adds complexity to data preparation.[59] The instruct versions of SahabatAI are tuned on a diverse set of instructions [13], and the sentiment task here might not align perfectly with those learned patterns.

The final choice between the Llama3-based or Gemma2-based SahabatAI variant will be made based on preliminary testing if time permits, or by considering factors like reported performance on closely related tasks in their respective benchmarks [11], and ease of use with available fine-tuning libraries. Both are strong candidates due to their recent architectures and extensive Indonesian pre-training.

3.4.2. Fine-Tuning Approach

- **Technique:** Parameter-Efficient Fine-Tuning (PEFT) will be the primary approach due to computational resource constraints typically faced in academic settings.
  - **LoRA (Low-Rank Adaptation) or QLoRA (Quantized Low-Rank Adaptation):** These methods significantly reduce the number of trainable parameters by injecting trainable low-rank matrices into the Transformer layers, while keeping the original pre-trained weights frozen.[92] QLoRA further reduces memory requirements by quantizing the pre-trained model to 4-bit precision.[60] This makes fine-tuning large models like SahabatAI (8B/9B parameters) feasible on single or limited GPU setups.
  - Full fine-tuning (updating all model parameters) will only be considered if PEFT methods yield unsatisfactory results and sufficient computational resources become available, which is unlikely for a Master's project.
- **Framework and Libraries:**
  - **Hugging Face Ecosystem:** The fine-tuning process will leverage the Hugging Face transformers library for model loading and basic operations, peft for implementing LoRA/QLoRA, datasets for data handling, and trl (Transformer Reinforcement Learning library, which also includes tools for Supervised Fine-Tuning like SFTTrainer) for managing the training loop.[91]
  - **Backend:** PyTorch will be the primary deep learning framework, given its widespread adoption and excellent support within the Hugging Face ecosystem. JAX is an alternative if specific advantages are identified.[60]

3.4.3. Hyperparameter Tuning

A systematic approach to hyperparameter tuning will be essential for optimizing the performance of the fine-tuned SahabatAI model. Key hyperparameters to be tuned include:

- **Learning Rate:** A small learning rate (e.g., in the range of 1×10−5 to 5×10−4) is typically used for fine-tuning LLMs.
- **Batch Size:** The number of training examples utilized in one iteration. This will be chosen based on GPU memory capacity (e.g., 4, 8, 16, 32).
- **Number of Epochs:** The number of times the entire training dataset is passed through the model (e.g., 1-5 epochs are common for fine-tuning).
- **Optimizer:** AdamW is a common choice for Transformer models.[30]
- **Weight Decay:** A regularization technique to prevent overfitting.
- **LoRA-specific parameters (if using LoRA/QLoRA):**
    - r (Rank of the decomposition): Determines the number of trainable parameters in LoRA layers (e.g., 8, 16, 32, 64).
    - lora_alpha (Scaling factor for LoRA).
    - target_modules: Specifying which modules (e.g., attention layers like q_proj, v_proj) in the Transformer architecture to apply LoRA to.[92] The validation set will be used to evaluate different hyperparameter configurations and select the combination that yields the best performance on the sentiment analysis task. The SahabatAI model cards provide some details on their own pre-training and instruction-tuning hyperparameters [11], which might offer useful starting points or ranges.

3.4.4. Computational Resources

- **GPU Requirements:** Fine-tuning an 8B or 9B parameter model, even with PEFT techniques like QLoRA, will require a GPU with substantial memory (e.g., NVIDIA A100 40GB/80GB, RTX 3090/4090 24GB, or V100 16GB/32GB). Access to university High-Performance Computing (HPC) clusters, or cloud-based GPU services (e.g., Google Colab Pro/Pro+, AWS SageMaker, GCP AI Platform, Azure ML) will be necessary. QLoRA is particularly designed to make fine-tuning feasible on consumer-grade GPUs with ~24GB VRAM, or even on Colab's T4 GPUs (16GB) for smaller ranks.[60]
- **Estimated Training Time:** Depending on the dataset size, batch size, number of epochs, and GPU used, fine-tuning could take several hours to a few days per experiment. The fine-tuning of SahabatAI's own instruct models took around 4 hours on 8x H100-80GB GPUs [13]; PEFT on a smaller dataset with a single GPU will be longer but manageable.
- **Software Environment:** A Python environment with up-to-date versions of PyTorch, Hugging Face libraries (transformers, peft, datasets, accelerate, trl, bitsandbytes for QLoRA), and other necessary NLP/data science packages (e.g., scikit-learn, pandas, numpy).

The choice of a SahabatAI base model, combined with PEFT techniques, represents a balanced approach to leveraging a powerful, Indonesian-specific LLM while managing the computational demands typically associated with such models.

3.5. Evaluation Framework

A comprehensive evaluation framework will be employed to assess the performance of the fine-tuned SahabatAI model on the Indonesian MoFA sentiment analysis task. This will involve quantitative metrics, baseline comparisons, and qualitative analysis.

3.5.1. Performance Metrics

The following standard classification metrics will be used to evaluate the model's performance on the held-out test set [33]:

- **Accuracy:** The proportion of correctly classified instances to the total number of instances. While common, it can be misleading for imbalanced datasets. Accuracy=TP+TN+FP+FNTP+TN (for binary, extendable to multi-class)
- **Precision (per class):** The proportion of correctly predicted positive instances (for a given class) to the total instances predicted as positive (for that class). Measures exactness. Precision=TP+FPTP
- **Recall (Sensitivity, per class):** The proportion of correctly predicted positive instances (for a given class) to the total actual positive instances (for that class). Measures completeness. Recall=TP+FNTP
- **F1-Score (per class, macro, micro, weighted):** The harmonic mean of precision and recall, providing a single score that balances both. F1–Score=2×Precision+RecallPrecision×Recall
    - **Macro F1:** Calculates F1 for each class and then takes the unweighted average. Treats all classes equally.
    - **Micro F1:** Calculates global TP, FP, FN and then the F1 score. More influenced by larger classes.
    - **Weighted F1:** Calculates F1 for each class and then takes the average, weighted by the number of true instances for each class. Accounts for class imbalance.
- **Confusion Matrix:** A table that visualizes the performance of the classification model, showing the counts of true positive, true negative, false positive, and false negative predictions for each class. Essential for understanding error patterns.[32]
- **Area Under the ROC Curve (AUC-ROC):** For binary classification tasks (or one-vs-rest for multi-class), this metric evaluates the model's ability to distinguish between classes across different thresholds.

Table 3.3: Evaluation Metrics for Sentiment Analysis

Metric	Formula / Description	Relevance to Sentiment Analysis Task
Accuracy	(TP+TN)/(TP+TN+FP+FN)	Overall correctness; can be misleading with imbalanced sentiment classes.
Precision (per class)	TP/(TP+FP)	For a given sentiment (e.g., Positive), how many texts classified as Positive were actually Positive? Important for minimizing false alarms.
Recall (per class)	TP/(TP+FN)	For a given sentiment (e.g., Negative), how many of the actual Negative texts did the model correctly identify? Important for capturing all instances of a sentiment.
F1-Score (per class, macro, weighted)	2×(Precision×Recall)/(Precision+Recall)	Balanced measure of precision and recall. Macro-F1 is good for overall performance across classes, especially if class balance is important. Weighted-F1 is useful for imbalanced datasets.
Confusion Matrix	Visual table of actual vs. predicted classes	Helps identify which sentiments are commonly confused with each other (e.g., Neutral vs. Positive).
AUC-ROC	Area under Receiver Operating Characteristic curve	Measures the model's ability to discriminate between sentiments across various thresholds; robust to class imbalance.

3.5.2. Baseline Comparisons

To provide context for the performance of the fine-tuned SahabatAI model, comparisons will be made against relevant baselines:

1. **SahabatAI Base Model (Zero-Shot/Few-Shot):** The chosen SahabatAI base model (Llama3 or Gemma2 variant) will be evaluated on the test set without any fine-tuning, using zero-shot or few-shot prompting techniques if applicable for classification with decoder-only models. This will establish the model's out-of-the-box capability for the task. However, for base models, direct classification is usually achieved via fine-tuning a classification head, so zero-shot prompting might be less effective than for instruct models.
2. **Fine-Tuned IndoBERT Model (Key Baseline):** A pre-trained IndoBERT model (e.g., indobenchmark/indobert-base-p1 or indobenchmark/indobert-large-p1 [10]) will be fine-tuned on the *exact same custom MoFA sentiment dataset* (training, validation, and test splits) used for SahabatAI. Models like taufiqdp/indonesian-sentiment [30] or crypter70/IndoBERT-Sentiment-Analysis [44] (which are already fine-tuned IndoBERT versions on general sentiment datasets like IndoNLU SmSA) could also be considered for further fine-tuning on the domain-specific data. This comparison is crucial for assessing whether SahabatAI offers tangible improvements over established Indonesian-specific models for this particular task.
3. **Traditional Machine Learning Model (Optional, if time permits):** A strong traditional ML model, such as Support Vector Machine (SVM) with TF-IDF features, could be trained and evaluated on the same dataset.[33] While deep learning models often outperform traditional methods on large datasets, SVMs can be competitive on smaller, text-based classification tasks and provide a non-neural baseline.[33]

The comparison against a domain-adapted IndoBERT is particularly important. It will help determine if SahabatAI's newer architecture and more extensive/recent Indonesian pre-training translate into superior performance on this highly specialized diplomatic sentiment analysis task, controlling for the fine-tuning dataset.

3.5.3. Qualitative Analysis

Beyond quantitative metrics, a qualitative analysis of the model's predictions will be conducted:

- **Error Analysis:** A sample of misclassified instances from the test set will be manually reviewed to identify patterns in errors. This can reveal specific linguistic constructions, topics, or sentiment expressions that the model struggles with (e.g., sarcasm, complex negations, highly nuanced diplomatic language).

- **Case Studies:** For selected MoFA statements or news articles related to significant events during Indonesia's ASEAN Chairmanship, the model's sentiment predictions will be examined in detail and compared with human interpretation and the known context of the events. This will help assess the real-world coherence and interpretability of the model's outputs.
- **Linguistic Challenges:** Observations on how the model handles specific linguistic features of Bahasa Indonesia relevant to diplomatic and media texts will be documented.

This multi-faceted evaluation approach, combining quantitative metrics, strong baselines, and qualitative insights, will provide a thorough assessment of the fine-tuned SahabatAI model's capabilities and limitations for sentiment analysis in the Indonesian MoFA domain.

3.6. Ethical Considerations in Data Handling and AI Application

The application of AI, particularly in sensitive domains like government communications and public opinion analysis, necessitates careful consideration of ethical implications. This research will adhere to the following ethical principles:

- **Data Privacy and Anonymization:**
  - **Official Communications and News Media:** Data from Kemlu RI's official website and mainstream news outlets are generally considered public information. However, care will be taken to ensure that any incidental personal data (e.g., names of non-public figures in comments sections if analyzed) is handled appropriately.
  - **Social Media Data (if used):** If Twitter/X data is collected, user identities (usernames, user IDs) will be anonymized or pseudonymized in any datasets used for training or public reporting to protect individual privacy, in line with ethical research practices for social media data. All data collection will comply with Twitter's terms of service and API usage policies.
  - **Compliance with Indonesian Regulations:** The research will be mindful of Indonesian data protection regulations (e.g., Law No. 27 of 2022 concerning Personal Data Protection) and AI ethics guidelines, such as those outlined in Indonesia's National Strategy for Artificial Intelligence (Stranas KA) and related ministerial circulars (e.g., Minister of Communication and Informatics Circular Letter No. 9 of 2023).[96] These frameworks emphasize principles like data governance, privacy, and lawful processing.

- **Bias Mitigation and Fairness:**
  - **Data Source Bias:** It is acknowledged that news media outlets may have inherent editorial slants, and social media discourse can be prone to echo chambers, polarization, and unrepresentative opinions. The selection of diverse news sources aims to mitigate some media bias, but the potential for such biases influencing the sentiment data will be discussed in the analysis of results.
  - **Algorithmic Bias:** LLMs like SahabatAI can inherit biases present in their vast pre-training data.[1] While SahabatAI's Indonesian-centric training may reduce certain biases compared to globally trained models, it is not immune. The research will be alert to potential biases (e.g., towards certain political viewpoints, entities, or demographic groups if reflected in the data) that could affect sentiment interpretation. The ASEAN Guide on AI Governance and Ethics, which Indonesia contributes to, also highlights the need for fair and inclusive AI.[100]
  - **Annotation Bias:** If manual annotation is performed, clear guidelines and potentially multiple annotators (for IAA on a subset) will be used to minimize subjective bias in labeling.[25]
- **Responsible Interpretation and Use of Results:**
  - The findings of the sentiment analysis will be presented with a clear acknowledgment of the methodology's limitations. Sentiment scores are interpretations and not absolute measures of truth or intent.
  - It will be emphasized that the developed tool is intended to provide insights and support human analysis, not to replace human judgment or be used for automated decision-making in critical foreign policy contexts.[1] This aligns with Indonesia's National AI Strategy, which calls for human oversight and accountability in AI systems.[96]
  - The potential for misuse of sentiment analysis tools in governmental contexts (e.g., for surveillance or manipulation of public opinion) is recognized. This research is framed as a contribution to enhancing understanding and improving public diplomacy through transparent means, not for control. The ethical guidelines within Indonesia's AI strategy, such as focusing on the benefit of humanity and accountability [101], will be guiding principles.
- **Transparency and Reproducibility:**
  - The research methodology, including data collection procedures, preprocessing steps, annotation guidelines (if applicable), fine-tuning parameters, and evaluation methods, will be documented in detail in the thesis.
  - To the extent permitted by data source terms of use and ethical

considerations (e.g., privacy), anonymized datasets and code developed for this research may be made available to the academic community to facilitate reproducibility and further research.

By proactively addressing these ethical considerations, this research aims to ensure that the development and application of AI for sentiment analysis in the Indonesian MoFA domain are conducted responsibly and contribute positively to both academic knowledge and practical understanding.

Chapter 4: Implementation Plan and Timeline

4.1. Detailed Work Plan (Six-Month Period)

The successful completion of this Master's thesis within a six-month timeframe requires a structured and realistic work plan. The major tasks and their estimated durations are outlined below and can be visualized in a Gantt chart format (descriptively here, to be visualized in the actual proposal as per [103]).

Table 4.1: Six-Month Research Timeline (Illustrative Gantt Chart)

Task ID	Task Description	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
		W1 W2 W3 W4	W1 W2 W3 W4	W1 W2 W3 W4	W1 W2 W3 W4	W1 W2 W3 W4	W1 W2 W3 W4
Phase 1: Foundation & Data Preparation (Months 1-2)							
T1.1	Literature Review Finalization & Refinement	XXXX					

T1.2	University Ethics Approval Process (if required for data handling)	XXXX	XX				
T1.3	Finalize Specific Research Focus (ASEAN Chairmanship 2023 details)	XX					
T1.4	Data Source Identification & API Access Confirmation/Application (NewsAPI, Twitter)		XXXX	XX			
T1.5	Development of Web Scraping & Data Collection Tools (Kemlu, News sites)		XX	XXXX	XX		
T1.6	Initial			XX	XXXX	XX	

	Data Collection (Kemlu, News, initial Twitter if API secured)						
Phase 2: Dataset Creation & Model Setup (Months 2-3)							
T2.1	Data Cleaning and Preprocessing Pipeline Development				XX	XXXX	
T2.2	Annotation Guideline Development & Pilot Annotation					XXXX	XX
T2.3	Main Data Annotation (Manual core set)						XXXX

	& Validation (IAA)						
T2.4	Sahabat AI & IndoBERT Environment Setup (Hugging Face, GPU access)				XX	XX	
Phase 3: Model Fine-Tuning & Evaluation (Months 4-5)							
T3.1	Fine-tuning Sahabat AI (Base Model) - Initial Experiments & Hyperparameter Tuning						XX
					XXXX	XX	
T3.2	Fine-tuning IndoBERT (Baselin					XXXX	XX

	e) - Experiments & Hyperparameter Tuning						
T3.3	Run Final Evaluation on Test Sets for all Models						XXXX
T3.4	Quantitative Results Analysis (Metrics, Comparisons)						
						XXXX	XX
Phase 4: Analysis, Writing & Submission (Months 5-6)							
T4.1	Qualitative Analysis (Error Analysis, Case Studies)						

						XX	XXXX
T4.2	Thesis Writing: Chapters 1-3 (Introduction, Lit Review, Methodology)		XX	XXXX	XX		
T4.3	Thesis Writing: Chapter 4 (Implementation, Timeline - update as actuals)					XX	
T4.4	Thesis Writing: Chapter 5 (Results, Discussion, Conclusion/Expected Outcomes)						XXXX
T4.5	Full Thesis Draft Compilation and Submission to Advisor						

							XXXX
T4.6	Revisions based on Advisor Feedback						
T4.7	Final Thesis Submission & Preparation for Defense						

Key: XXXX represents full month activity, XX represents partial month activity.

- **Detailed Activities & Milestones:**
  - **Month 1:** Complete literature review, secure ethics approval if needed for handling any potentially sensitive data aspects (even if public), finalize the precise scope of the ASEAN Chairmanship 2023 focus. Begin development of scraping tools. Initiate API access requests. *Milestone: Finalized research scope, ethics clearance (if applicable).*
  - **Month 2:** Complete development of scraping tools. Begin data collection from Kemlu RI website and selected news outlets. Continue pursuing API access. Begin developing the data cleaning and preprocessing pipeline. *Milestone: Initial raw data corpus collected from Kemlu and at least two news sources. Preprocessing pipeline functional.*
  - **Month 3:** Complete primary data collection. Finalize and apply data cleaning/preprocessing. Develop annotation guidelines. Conduct pilot annotation and refine guidelines. Begin main manual annotation of the core dataset. Set up the computational environment for SahabatAI and IndoBERT. *Milestone: Cleaned and preprocessed dataset ready. Annotation guidelines finalized. Annotation of core set initiated. GPU environment tested.*

- **Month 4:** Complete manual annotation of the core dataset and calculate IAA. Split data into train/validation/test sets. Begin fine-tuning experiments for SahabatAI (base model), focusing on hyperparameter optimization using the validation set. *Milestone: Fully annotated and split dataset. Initial fine-tuning results for SahabatAI.*
- **Month 5:** Complete SahabatAI fine-tuning and evaluation. Fine-tune IndoBERT model on the same dataset as a baseline. Run final evaluations for all models on the test set. Compile and analyze quantitative results. Begin qualitative error analysis. Continue thesis writing (Chapters 1-3 should be well underway). *Milestone: All model evaluations complete. Quantitative results analyzed.*
- **Month 6:** Complete qualitative analysis. Finalize results and discussion. Complete writing of all thesis chapters. Submit full draft to advisor(s). Incorporate feedback and prepare final thesis for submission and defense. *Milestone: Full thesis draft submitted. Final thesis submitted after revisions.*

4.2. Required Resources

- **Computational Resources:**
  - **GPU Access:** Essential for fine-tuning SahabatAI and IndoBERT. Access to at least one GPU with >=16GB VRAM (e.g., NVIDIA V100, RTX 3080/3090/4090, A4000) is required. QLoRA can make 8B/9B models trainable on such hardware.[60] Access will be sought through university HPC facilities or services like Google Colab Pro/Pro+. Estimated GPU time: 50-100 hours for experiments, hyperparameter tuning, and final model training runs, depending on the efficiency of PEFT and the number of experiments.
- **Software and Libraries:**
  - **Programming Language:** Python 3.8+
  - **Core Libraries:** PyTorch, Hugging Face Suite (transformers, datasets, tokenizers, peft, accelerate, trl) [91], scikit-learn (for metrics, traditional ML baseline), pandas (data manipulation), numpy (numerical operations).
  - **Web Scraping:** BeautifulSoup4, requests, Scrapy (if needed).
  - **PDF Text Extraction:** PyPDF2, pdfminer.six, or similar.
  - **Annotation Tool (if dedicated tool used beyond spreadsheets):** Label Studio, Doccano (open-source options), or custom scripts for managing annotation.
- **Data Access and Storage:**
  - **API Keys/Subscriptions:** Potential costs for NewsAPI.org or Twitter API (Basic Tier at $200/month [22] if used for a limited period). Budgetary approval or alternative strategies will be needed if these costs are significant.

    - **Storage:** Approximately 50-100GB of local or cloud storage for raw data, preprocessed data, annotated datasets, model checkpoints, and experimental results.
- **Human Resources:**
    - **Primary Researcher Time:** Significant time allocation for all phases, especially data collection, cleaning, annotation (if manual), model experimentation, and thesis writing.
    - **Annotators (if applicable beyond researcher):** If additional annotators are involved for IAA, they would need to be native Indonesian speakers familiar with the annotation guidelines. This is likely limited to the researcher and perhaps one other for a pilot study in a Master's context.

Early confirmation of GPU access and a clear plan for API key acquisition (or alternative data collection strategies) are critical for adhering to the timeline.

4.3. Risk Assessment and Mitigation Strategies

Several potential risks could impact the successful completion of this research. Proactive identification and mitigation strategies are outlined below:

- **Technical Risks:**
    - **Difficulties in Fine-tuning SahabatAI:** As SahabatAI is relatively new, unforeseen challenges in stability, convergence, or achieving desired performance during fine-tuning might arise.
        - **Mitigation:** Start with recommended hyperparameters from SahabatAI documentation (if available from model cards [11]) or similar Llama3/Gemma2 fine-tuning guides.[92] Systematically experiment with learning rates, batch sizes, and LoRA configurations. Consult Hugging Face community forums and SahabatAI's collaboration channels [11] if specific issues arise. Ensure the software environment is correctly configured.
    - **Challenges in Indonesian Text Preprocessing:** Handling diverse linguistic phenomena in Bahasa Indonesia (e.g., novel slang, complex morphological variations not covered by Sastrawi, inconsistent spelling in informal texts) could be more complex than anticipated.
        - **Mitigation:** Employ an iterative approach to preprocessing. Start with standard techniques and refine the pipeline based on observed issues in the data. Augment slang dictionaries as new terms are encountered. Use robust tokenizers provided with SahabatAI models.
- **Data-related Risks:**
    - **Insufficient Data Volume or Quality:** The collected data from MoFA, news,

or social media might be less voluminous or of lower quality (e.g., too much noise, irrelevant content) than expected for the chosen ASEAN Chairmanship focus.
    - **Mitigation:** Broaden search queries slightly if initial collection is too narrow. Be prepared to include data from a wider range of news outlets if the primary ones are insufficient (though this adds to scraping complexity). If necessary, slightly adjust the research focus to match the available data, or extend the data collection period if the timeline permits. Prioritize data quality over quantity for the annotation phase.
  - **Low Inter-Annotator Agreement (IAA):** If manual annotation is performed by multiple annotators (even for a pilot), achieving satisfactory IAA can be challenging for a subjective task like sentiment analysis, especially with nuanced diplomatic text.
    - **Mitigation:** Develop highly detailed and clear annotation guidelines with numerous examples (as per Table 3.2). Conduct thorough training sessions for annotators. Perform multiple rounds of pilot annotation and guideline refinement based on disagreements. Implement a clear adjudication process for resolving conflicting labels.
  - **API Access Changes or Restrictions:** Terms of service, pricing, or availability of NewsAPI.org or the Twitter API can change with little notice.[20]
    - **Mitigation:** For news data, have web scraping scripts as a backup for all targeted outlets. For Twitter data, if API access becomes unfeasible, the scope of social media analysis will be significantly reduced or removed, with a focus shifted more heavily onto MoFA and news media data. Clearly state this dependency in the research.
- **Time-related Risks:**
  - **Underestimation of Time for Key Tasks:** Data collection (especially if relying on extensive scraping), manual data annotation, and iterative model experimentation can be very time-consuming.
    - **Mitigation:** Build some buffer time into the Gantt chart for critical tasks. Prioritize the core research objectives (fine-tuning SahabatAI on MoFA/news data). If significant delays occur, be prepared to reduce the scope (e.g., limit the number of news sources, reduce the size of the annotated dataset while focusing on quality, or drop the Twitter analysis component if it proves too time-consuming or costly to access data). Maintain regular progress reviews with the thesis advisor.

The dynamic nature of APIs and the rapid evolution of LLMs and their associated libraries mean that flexibility and adaptability will be key throughout the research process. Continuous monitoring of API terms and software updates will be necessary.

Chapter 5: Expected Outcomes and Dissemination

5.1. Reiteration of Anticipated Contributions

This research is expected to yield several significant contributions, spanning both academic advancements in Indonesian NLP and practical insights for the Indonesian Ministry of Foreign Affairs (Kemlu RI).

- **Academic Contributions:**
    - **Benchmarking SahabatAI:** The core academic outcome will be a rigorous evaluation and benchmark of a fine-tuned SahabatAI model (Llama3 or Gemma2 variant) for sentiment analysis in the specialized domain of Indonesian foreign policy and diplomatic communications. This will be one of the early in-depth studies applying SahabatAI to a complex, nuanced, and low-resource task beyond general NLP benchmarks, providing valuable data on its adaptability and performance.
    - **Contribution to Indonesian NLP Resources:** Should manual annotation be extensively performed, the resulting domain-specific annotated dataset (focused on Kemlu RI's ASEAN Chairmanship communications and related media) will be a valuable addition to Indonesian NLP resources, potentially made available to the research community (with appropriate anonymization and adherence to data use terms).
    - **Insights into Indonesian Diplomatic Language Analysis:** The research will shed light on the specific linguistic challenges (e.g., handling officialese, detecting subtle sentiment in formal statements, processing media narratives about diplomacy) and effective methodologies for sentiment analysis of political and diplomatic texts in Bahasa Indonesia.
- **Practical Contributions for Kemlu RI:**
    - **Tool for Kemlu RI:** The fine-tuned SahabatAI model has the potential to serve as a prototype or foundational component for a tool that Kemlu RI could use for enhanced media monitoring, analysis of public discourse surrounding its policies, and understanding the reception of its strategic communications. This aligns with the increasing need for data-driven insights in modern diplomacy.[1]
    - **Understanding Policy Perception:** The analysis will offer a data-driven perspective on how Kemlu RI's communications and Indonesia's role during the 2023 ASEAN Chairmanship were perceived in selected Indonesian media, providing nuanced insights beyond anecdotal evidence.

This work aims to demonstrate the tangible benefits of applying advanced, locally-focused AI models like SahabatAI to address specific national interests, contributing to the broader goals of Indonesia's National AI Strategy [96] and its efforts in digital diplomacy.[6]

5.2. Potential for Future Research

This Master's thesis can serve as a foundational study, opening several avenues for future research:

- **Expansion of Sentiment Analysis:**
  - **Granular Emotion Analysis:** Moving beyond positive/negative/neutral to detect more specific emotions (e.g., trust, anger, anticipation, surprise) in diplomatic texts using models like SahabatAI.[25]
  - **Aspect-Based Sentiment Analysis (ABSA):** A more detailed investigation into ABSA to determine sentiment towards specific facets or themes within Indonesian foreign policy (e.g., sentiment towards economic aspects vs. political aspects of a policy).[8]
- **Broader NLP Applications in Diplomacy:**
  - **Topic Modeling:** Identifying key themes and shifts in focus within Kemlu RI communications or media coverage over time.
  - **Argument Mining:** Extracting and analyzing the structure of arguments presented in diplomatic statements or debates.
  - **Detection of Misinformation/Disinformation and Propaganda:** Applying NLP techniques to identify manipulative narratives related to Indonesian foreign policy, a growing concern in international relations.[51]
  - **Explainable AI (XAI) for Diplomatic Insights:** Developing and applying XAI techniques to understand *why* the SahabatAI model makes certain sentiment predictions, which is crucial for building trust and utility in sensitive domains like diplomacy.
- **Cross-Lingual and Multimodal Analysis:**
  - Extending the analysis to include sentiment in other languages relevant to Indonesian diplomacy (e.g., English-language international media).
  - Incorporating multimodal data, such as analyzing images or videos that accompany news articles or social media posts related to foreign policy events, for a richer understanding of the communicated message and its sentiment.
- **Longitudinal and Comparative Studies:**
  - Conducting longitudinal studies to track the evolution of sentiment towards Indonesian foreign policy over extended periods.

    - Performing comparative sentiment analysis across different countries' MoFAs or on similar policy issues as addressed by various nations.
- **Integration with Other AI Tools for Diplomacy:**
    - Exploring how sentiment analysis outputs can be integrated with other AI-driven diplomatic tools, such as those for early warning systems for conflict [50], negotiation support, or simulating diplomatic scenarios.[50]

The successful application of SahabatAI in this thesis can act as a catalyst for broader exploration and adoption of AI technologies within Kemlu RI and the wider Indonesian public sector. This aligns with Indonesia's national vision for leveraging AI for societal benefit and enhancing governance [109], and can contribute to building a robust ecosystem for "AI for Diplomacy" in Indonesia.

5.3. Dissemination Plan

The findings of this research will be disseminated through the following channels:

1. **Master's Thesis:** The primary output will be the submission of the completed Master's Thesis to the [User's University] in fulfillment of degree requirements.
2. **Academic Publication (Potential):** Subject to the quality and novelty of the findings, a paper will be prepared for submission to a relevant peer-reviewed academic journal or conference. Potential venues include:
    - NLP conferences (e.g., ACL, EMNLP, AACL-IJCNLP, or regional workshops focusing on Southeast Asian languages).
    - Journals focusing on computational linguistics, AI applications, political science with a computational methodology, or Indonesian/Southeast Asian studies. The Indonesian Journal of International Relations (IJIR) [111] could be a relevant local venue.
3. **Open Science Contributions (Conditional):**
    - **Dataset:** If a new, manually annotated dataset is created, and subject to ethical considerations and data source terms of use, an anonymized version of this dataset may be shared with the research community via platforms like Hugging Face Datasets or a university repository. This would contribute valuable resources for Indonesian NLP.
    - **Code:** Key scripts developed for data preprocessing, model fine-tuning (especially PEFT configurations for SahabatAI), and evaluation may be shared via GitHub to promote reproducibility and further research.
4. **Presentation (Potential):** Findings may be presented at academic seminars, workshops, or student research symposiums.

This dissemination plan aims to maximize the impact of the research, contributing to academic knowledge and potentially informing practical applications in the field of diplomacy and AI.

References

(A comprehensive list of all cited sources will be compiled here, following a consistent academic citation style, e.g., APA 7th Edition or IEEE. For brevity in this proposal, only snippet IDs have been used. The final thesis will contain full citations for all sources like [68], etc.)

Works cited

1. AI-Powered Diplomacy: The Role of Artificial Intelligence in Global Conflict Resolution, accessed May 10, 2025, https://trendsresearch.org/insight/ai-powered-diplomacy-the-role-of-artificial-intelligence-in-global-conflict-resolution/
2. AI in Diplomacy: How Technology is Transforming International Relations - The Diplomatist, accessed May 10, 2025, https://diplomatist.com/2025/02/14/ai-in-diplomacy-how-technology-is-transforming-international-relations/
3. The meaning of sentiment analysis of UN speeches on the Russia-Ukraine war: a comparative study using VADER and BERT NLP techniques - Frontiers, accessed May 10, 2025, https://www.frontiersin.org/articles/10.3389/fpos.2025.1546822/full
4. (PDF) The meaning of sentiment analysis of UN speeches on the Russia-Ukraine war: a comparative study using VADER and BERT NLP techniques - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/390506430_The_meaning_of_sentiment_analysis_of_UN_speeches_on_the_Russia-Ukraine_war_a_comparative_study_using_VADER_and_BERT_NLP_techniques
5. Introduction to Sentiment Analysis in natural language processing (NLP) - Project Guru, accessed May 10, 2025, https://www.projectguru.in/introduction-to-sentiment-analysis-in-natural-language-processing-nlp/
6. e-journal.unair.ac.id, accessed May 10, 2025, https://e-journal.unair.ac.id/JHI/article/download/17996/10686/73865
7. PENTINGNYA DIPLOMASI DIGITAL DALAM PENCAPAIAN ..., accessed May 10, 2025, https://pacis.unpar.ac.id/pentingnya-diplomasi-digital-dalam-pencapaian-kebijakan-luar-negeri/
8. Aspect-based Sentiment Analysis and Topic Modelling of International Media on Indonesia Tourism Sector Recovery - Lasigo Journals, accessed May 10, 2025, https://journal.lasigo.org/index.php/IJTL/article/download/502/219
9. arxiv.org, accessed May 10, 2025, https://arxiv.org/pdf/2403.01817
10. [2009.05387] IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding - ar5iv, accessed May 10, 2025,

https://ar5iv.labs.arxiv.org/html/2009.05387
11. GoToCompany/llama3-8b-cpt-sahabatai-v1-base · Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/llama3-8b-cpt-sahabatai-v1-base
12. GoToCompany/gemma2-9b-cpt-sahabatai-v1-base · Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/gemma2-9b-cpt-sahabatai-v1-base
13. GoToCompany/llama3-8b-cpt-sahabatai-v1-instruct · Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/llama3-8b-cpt-sahabatai-v1-instruct
14. Sahabat-AI, accessed May 10, 2025, https://sahabat-ai.com/
15. Llama3 8b Cpt Sahabatai V1 Instruct · Models - Dataloop AI, accessed May 10, 2025, https://dataloop.ai/library/model/gotocompany_llama3-8b-cpt-sahabatai-v1-instruct/
16. Mapping Sentiments: A Journey into Low-Resource Luxembourgish Analysis - ACL Anthology, accessed May 10, 2025, https://aclanthology.org/2024.luhme-1.3.pdf
17. Explainable Pre-Trained Language Models for Sentiment Analysis in Low-Resourced Languages - MDPI, accessed May 10, 2025, https://www.mdpi.com/2504-2289/8/11/160
18. Sarcasm Detection For Sentiment Analysis in Indonesian Tweets - CORE, accessed May 10, 2025, https://core.ac.uk/works/83065941
19. Sarcasm Detection For Sentiment Analysis in Indonesian Tweets | Yunitasari | IJCCS (Indonesian Journal of Computing and Cybernetics Systems), accessed May 10, 2025, https://jurnal.ugm.ac.id/ijccs/article/view/41136
20. Twitter API Alternative: Building Applications After API Restrictions - The Suffolk News-Herald, accessed May 10, 2025, https://www.suffolknewsherald.com/2025/04/24/twitter-api-alternative-building-applications-after-api-restrictions/
21. Twitter API v2 or Third-party Tool: Which One Is Right for You? | Data365.co, accessed May 10, 2025, https://data365.co/blog/twitter-api-v2
22. About the X API - X, accessed May 10, 2025, https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api
23. How to register Twitter API for Academic Research (Free), accessed May 10, 2025, https://nariyoo.com/how-to-register-academic-twitter-api/
24. Perception of the Indonesian Society on the Performance of the Ministry of Investment/BKPM Based on Sentiment Analysis - International Journal of Engineering Trends and Technology, accessed May 10, 2025, https://ijettjournal.org/Volume-72/Issue-2/IJETT-V72I2P125.pdf
25. Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm - MDPI, accessed May 10, 2025, https://www.mdpi.com/2078-2489/15/11/705
26. (PDF) Exploring emotion classification of indonesian tweets using large scale

transfer learning via IndoBERT - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/390029900_Exploring_emotion_classification_of_indonesian_tweets_using_large_scale_transfer_learning_via_IndoBERT
27. SENTIMENT ANALYSIS ON INDONESIAN SOCIAL MEDIA USING INTEGRATION OF INSET-LEXICON AND MACHINE LEARNING - Repository Nusa Putra University, accessed May 10, 2025, https://repository.nusaputra.ac.id/1382/1/NOVA%20TEGUH%20KRISDWIYANTO%20.pdf
28. An experimental study of lexicon-based sentiment analysis on Bahasa Indonesia, accessed May 10, 2025, https://www.researchgate.net/publication/312560941_An_experimental_study_of_lexicon-based_sentiment_analysis_on_Bahasa_Indonesia
29. A Multi-Labeled Dataset for Indonesian Discourse: Examining Toxicity, Polarization, and Demographics Information - arXiv, accessed May 10, 2025, https://arxiv.org/html/2503.00417v1
30. taufiqdp/indonesian-sentiment - Hugging Face, accessed May 10, 2025, https://huggingface.co/taufiqdp/indonesian-sentiment
31. Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/portal/id
32. Sentiment analysis based on Indonesian language lexicon and IndoBERT on user reviews PLN mobile application - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/390378101_Sentiment_analysis_based_on_Indonesian_language_lexicon_and_IndoBERT_on_user_reviews_PLN_mobile_application
33. Sentiment Analysis on a Large Indonesian Product Review Dataset - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/378755917_Sentiment_Analysis_on_a_Large_Indonesian_Product_Review_Dataset
34. A comparative study of sentiment analysis using SVM and Senti Word Net - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/332141053_A_comparative_study_of_sentiment_analysis_using_SVM_and_Senti_Word_Net
35. Sentiment Analysis of the Public Towards the Kanjuruhan Tragedy with the Support Vector Machine Method - Jurnal UMP, accessed May 10, 2025, https://jurnalnasional.ump.ac.id/index.php/JUITA/article/download/17405/6449
36. Sentiment analysis of Indonesian datasets based on a hybrid deep ..., accessed May 10, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC10226016/
37. COMPARISON OF PRE-TRAINED BERT-BASED TRANSFORMER MODELS FOR REGIONAL LANGUAGE TEXT SENTIMENT ANALYSIS IN INDONESIA - Semantic Scholar, accessed May 10, 2025, https://pdfs.semanticscholar.org/0e95/b4b6dcdfd9952da9cfa7da512b88c6600499.pdf
38. Sentiment Analysis on a Large Indonesian Product Review Dataset | Journal of Information Systems Engineering and Business Intelligence, accessed May 10, 2025, https://e-journal.unair.ac.id/JISEBI/article/view/48215
39. NusaCrowd: Open Source Initiative for Indonesian NLP Resources - ACL

Anthology, accessed May 10, 2025, https://aclanthology.org/2023.findings-acl.868v2.pdf
40. Sentiment Analysis for Low Resource Languages: A Study on Informal Indonesian Tweets, accessed May 10, 2025, https://www.researchgate.net/publication/320944752_Sentiment_Analysis_for_Low_Resource_Languages_A_Study_on_Informal_Indonesian_Tweets
41. Multilingual Propaganda Detection: Exploring Transformer-Based Models mBERT, XLM-RoBERTa, and mT5 - ACL Anthology, accessed May 10, 2025, https://aclanthology.org/2025.nakbanlp-1.9.pdf
42. Master's Thesis Proposal Tips - University of Idaho, accessed May 10, 2025, https://www.uidaho.edu/cogs/resources/student-resources/thesis-dissertation/proposal-tips
43. SmSA (IndoNLU) Dataset - NLP Hub - Metatext - AI, accessed May 10, 2025, https://metatext.io/datasets/smsa-(indonlu)
44. crypter70 / IndoBERT-Sentiment-Analysis, accessed May 10, 2025, https://www.toolify.ai/ai-model/crypter70-indobert-sentiment-analysis
45. Sentiment-Analysis-with-IndoBERT-Fine-tuning-and-IndoNLU-SmSA-Dataset/LICENSE at main - GitHub, accessed May 10, 2025, https://github.com/crypter70/Sentiment-Analysis-with-IndoBERT-Fine-tuning-and-IndoNLU-SmSA-Dataset/blob/main/LICENSE
46. GoToCompany/gemma2-9b-cpt-sahabatai-v1-instruct · Hugging Face, accessed May 10, 2025, https://huggingface.co/GoToCompany/gemma2-9b-cpt-sahabatai-v1-instruct
47. www.diplomacy.edu, accessed May 10, 2025, https://www.diplomacy.edu/topics/ai-and-diplomacy/#:~:text=AI%20as%20a%20tool%20for%20diplomacy,-AI%20is%20not&text=AI%20technologies%20are%20being%20leveraged,in%20formulating%20informed%20policy%20decisions.
48. AI DIPLOMACY: geo-politics, topics and tools in 2025 | Diplo, accessed May 10, 2025, https://www.diplomacy.edu/topics/ai-and-diplomacy/
49. AI's Transformative Impact On Diplomacy: A New Era Of Global Engagement – OpEd, accessed May 10, 2025, https://www.eurasiareview.com/30042025-ais-transformative-impact-on-diplomacy-a-new-era-of-global-engagement-oped/
50. (PDF) AI DIPLOMACY Insights and Innovations from the Bilateral ..., accessed May 10, 2025, https://www.researchgate.net/publication/390243157_AI_DIPLOMACY_Insights_and_Innovations_from_the_Bilateral_Navigator
51. www.defensa.gob.es, accessed May 10, 2025, https://www.defensa.gob.es/documents/2073105/2320762/analisis_de_la_geopolitica_mundial_mediante_inteligencia_artificial_ia_y_big_data_2025_dieeea01_eng.pdf
52. AI Use Cases for Diplomats: Applying Artificial Intelligence to Diplom - Routledge, accessed May 10, 2025, https://www.routledge.com/AI-Use-Cases-for-Diplomats-Applying-Artificial-Intelligence-to-Diplomacy/Kilburg/p/book/9781041008354

53. Political-LLM: Large Language Models in Political Science, accessed May 10, 2025, https://political-llm.org/
54. Full article: China's Discursive Strategic Neutrality Position in the ..., accessed May 10, 2025, https://www.tandfonline.com/doi/full/10.1080/10670564.2025.2479019?src=exp-la
55. ANALISIS FRAMING: PEMBERITAAN MEDIA MASSA MENGENAI KEBIJAKAN LUAR NEGERI INDONESIA DALAM MENANGGAPI RIVALITAS AMERIKA SERIKAT DA, accessed May 10, 2025, https://ejournal.fisip.unjani.ac.id/index.php/jurnal-dinamika-global/article/download/1070/357
56. ANALISIS FRAMING: PEMBERITAAN MEDIA MASSA MENGENAI KEBIJAKAN LUAR NEGERI INDONESIA DALAM MENANGGAPI RIVALITAS AMERIKA SERIKAT DAN CHINA (2019-2022) | Dinamika Global : Jurnal Ilmu Hubungan Internasional, accessed May 10, 2025, https://ejournal.fisip.unjani.ac.id/index.php/jurnal-dinamika-global/article/view/1070
57. irfnrdh/Awesome-Indonesia-NLP: Resource NLP & Bahasa - GitHub, accessed May 10, 2025, https://github.com/irfnrdh/Awesome-Indonesia-NLP
58. ailabtelkom/id-NLP-resources: Kumpulan resource untuk pemrosesan bahasa alami Bahasa Indonesia. Segala bentuk kontribusi sangat "WELCOME" - GitHub, accessed May 10, 2025, https://github.com/ailabtelkom/id-NLP-resources
59. Base LLM vs Instruction-Tuned LLM - Analytics Vidhya, accessed May 10, 2025, https://www.analyticsvidhya.com/blog/2025/02/base-llm-vs-instruction-tuned-llm/
60. Fine-Tune Gemma using Hugging Face Transformers and QloRA | Google AI for Developers, accessed May 10, 2025, https://ai.google.dev/gemma/docs/core/huggingface_text_finetune_qlora
61. Dataset Indonesia untuk Analisis Sentimen - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/profile/Wiliam-Fajar/publication/338409000_Dataset_Indonesia_untuk_Analisis_Sentimen/links/5e30d7e992851c7f7f08c002/Dataset-Indonesia-untuk-Analisis-Sentimen.pdf
62. diplomasi publik indonesia melalui bali democracy forum - ResearchGate, accessed May 10, 2025, https://www.researchgate.net/publication/356908555_DIPLOMASI_PUBLIK_INDONESIA_MELALUI_BALI_DEMOCRACY_FORUM
63. DIPLOMASI PUBLIK INDONESIA MELALUI BALI DEMOCRACY FORUM - Ejournal FISIP Unjani, accessed May 10, 2025, https://ejournal.fisip.unjani.ac.id/index.php/jurnal-dinamika-global/article/download/813/281
64. SIMPLE YET EFFECTIVE: RELATIONAL PUBLIC DIPLOMACY BY INDONESIAN ART CENTRES, accessed May 10, 2025, https://ijaps.usm.my/wp-content/uploads/2025/02/IJAPS-211_ART-5.pdf
65. THE IMPLEMENTATION OF DARMASISWA SCHOLARSHIP AS THE MEANS OF INDONESIA PUBLIC DILOMACY IN 2011 TO 2019 UNDERGRADUATE THESIS Writ,

accessed May 10, 2025, https://dspace.uii.ac.id/bitstream/handle/123456789/49282/19323123.pdf?sequence=1&isAllowed=y
66. Base LLM vs. instruction-tuned LLM - Toloka, accessed May 10, 2025, https://toloka.ai/blog/base-llm-vs-instruction-tuned-llm/
67. accessed January 1, 1970, https://kemlu.go.id/portal/id/search?searchtype=siaran-pers
68. Siaran Pers - Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/publikasi/siaran-pers
69. MENTERI LUAR NEGERI REPUBLIK INDONESIA PERATURAN MENTERI LUAR NEGERI REPUBLIK INDONESIA NOMOR 2 TAHUN 2024 TENTANG PERUBAHAN AT - Peraturan BPK, accessed May 10, 2025, https://peraturan.bpk.go.id/Download/335484/Permenlu%20Nomor%202%20Tahun%202024.pdf
70. Pernyataan Pers Tahunan Menteri Luar Negeri (PPTM) - Melbourne, accessed May 10, 2025, https://kemlu.go.id/melbourne/publikasi/pernyataan-pers-tahunan-menteri-luar-negeri-pptm
71. SUMBER-SUMBER ARSIP (GUIDE ARSIP) PIDATO PRESIDEN SUKARNO DI SIDANG UMUM PBB KE- 15 NEW YORK, 30 SEPTEMBER 1960 - ANRI, accessed May 10, 2025, https://anri.go.id/download/sumbersumber-arsip-guide-arsip-pidato-presiden-sukarno-di-sidang-umum-pbb-ke15-new-york-30-september-1960-1640069491
72. Kemlu Perkuat Infrastruktur Hukum dalam Selesaikan Masalah WNI di Luar Negeri, accessed May 10, 2025, https://www.hukumonline.com/berita/a/kemlu-perkuat-infrastruktur-hukum-dalam-selesaikan-masalah-wni-di-luar-negeri-lt66f201dc033e8/
73. jurnal - hubungan luar negeri - Portal Kemlu, accessed May 10, 2025, https://kemlu.go.id/files/repositori/56596/172775678566fb79f1e67da_JurnalHubluVol9No1JanuariJuni2024_001.pdf
74. Api - ANTARA News, accessed May 10, 2025, https://en.antaranews.com/tag/api
75. Everything - Documentation - News API, accessed May 10, 2025, https://newsapi.org/docs/endpoints/everything
76. Indonesia News API - Live top headlines from Indonesia, accessed May 10, 2025, https://newsapi.org/s/indonesia-news-api
77. API Untuk Mendapatkan Konten “KOMPAS.ID” – Knowledge Base ..., accessed May 10, 2025, https://kb.kompas.id/baca/partnership/api-untuk-mendapatkan-content-kompas-id/
78. (PDF) Transformasi Harian Kompas MenjadiPortal Berita Digita Subscription Kompas.Id, accessed May 10, 2025, https://www.researchgate.net/publication/335003164_Transformasi_Harian_Kompas_MenjadiPortal_Berita_Digita_Subscription_KompasId
79. Berita Kebijakan Luar Negeri Terbaru - Tempo.co, accessed May 10, 2025, https://www.tempo.co/tag/kebijakan-luar-negeri

80. Syarat dan Cara Menukar Poin di Access by KAI untuk Pembayaran Tiket Kereta Api, accessed May 10, 2025, https://www.tempo.co/digital/syarat-dan-cara-menukar-poin-di-access-by-kai-untuk-pembayaran-tiket-kereta-api-1394308
81. Berapa Biaya Pembatalan Tiket Kereta Api? Simak Ketentuannya di Sini - detikNews, accessed May 10, 2025, https://news.detik.com/berita/d-7201333/berapa-biaya-pembatalan-tiket-kereta-api-simak-ketentuannya-di-sini
82. Jenis, Syarat dan Cara untuk Mendapatkan Tarif Reduksi Tiket KAI - detikNews, accessed May 10, 2025, https://news.detik.com/berita/d-7014304/jenis-syarat-dan-cara-untuk-mendapatkan-tarif-reduksi-tiket-kai
83. Live top entertainment headlines from Indonesia - News API, accessed May 10, 2025, https://newsapi.org/s/indonesia-entertainment-news-api
84. accessed January 1, 1970, https://newsapi.org/docs/sources
85. Usage policies - OpenAI, accessed May 10, 2025, https://openai.com/policies/usage-policies/
86. (2025) Twitter API Intro: A Guide to Use and Test It Online - Apidog, accessed May 10, 2025, https://apidog.com/blog/what-is-twitter-api-2024-guide/
87. accessed January 1, 1970, https://developer.twitter.com/en/products/twitter-api/academic-research
88. xdevplatform/getting-started-with-the-twitter-api-v2-for-academic-research - GitHub, accessed May 10, 2025, https://github.com/xdevplatform/getting-started-with-the-twitter-api-v2-for-academic-research
89. ANALISIS SENTIMEN DATA TWITTER TERHADAP PELAKSANAAN PEMBELAJARAN ONLINE DI INDONESIA PADA MASA PANDEMI COVID-19 MENGGUNAKAN METO, accessed May 10, 2025, https://jursistekni.nusaputra.ac.id/article/download/205/105
90. Pengambilan Data Twitter (X): Alat dan Praktik Terbaik 2025 - RapidSeedbox, accessed May 10, 2025, https://www.rapidseedbox.com/id/blog/mastering-twitter-scraping
91. Fine-Tuning LLMs: A Guide With Examples - DataCamp, accessed May 10, 2025, https://www.datacamp.com/tutorial/fine-tuning-large-language-models
92. Fine-Tuning 1B LLaMA 3.2: A Comprehensive Step-by-Step Guide with Code, accessed May 10, 2025, https://huggingface.co/blog/ImranzamanML/fine-tuning-1b-llama-32-a-comprehensive-article
93. Fine-tune Gemma in Keras using LoRA | Google AI for Developers - Gemini API, accessed May 10, 2025, https://ai.google.dev/gemma/docs/core/lora_tuning
94. Fine-Tuning Llama 3.1 for Text Classification - DataCamp, accessed May 10, 2025, https://www.datacamp.com/tutorial/fine-tuning-llama-3-1
95. LLM Evaluation: Key Metrics, Best Practices and Frameworks - Aisera, accessed May 10, 2025, https://aisera.com/blog/llm-evaluation/
96. Indonesia | Global AI Ethics and Governance Observatory - UNESCO, accessed

May 10, 2025, https://www.unesco.org/ethics-ai/en/indonesia
97. Transforming Government Decision-Making by Utilizing Artificial Intelligence (AI) in the Future - Open Journal Systems, accessed May 10, 2025, https://jurnal.iapa.or.id/proceedings/article/download/1206/583
98. Indonesia | Herbert Smith Freehills | Global law firm, accessed May 10, 2025, https://www.herbertsmithfreehills.com/insights/reports/searchlight-ai/indonesia
99. Scale AI & CSIS Introduce Foreign Policy Decision Benchmark, accessed May 10, 2025, https://scale.com/blog/scale-csis-foreign-policy-benchmark
100. Advancing Safe, Trustworthy and Ethical AI for Climate Action - Arup, accessed May 10, 2025, https://www.arup.com/globalassets/downloads/insights/a/ai-for-sustainable-development-in-asean/ai-for-sustainable-development-in-asean.pdf
101. Artificial intelligence in public administration: benefits and risks - Dialnet, accessed May 10, 2025, https://dialnet.unirioja.es/descarga/articulo/9953186.pdf
102. Integrating Artificial Intelligence into Public Administration: Challenges and Vulnerabilities, accessed May 10, 2025, https://www.mdpi.com/2076-3387/15/4/149
103. Format for a Thesis Proposal - Cal Poly Pomona, accessed May 10, 2025, https://www.cpp.edu/sci/computer-science/current-graduate-students/handbook/proposal-form.shtml
104. Indonesia's AI Strategy: Balancing AI Regulation and Innovation ..., accessed May 10, 2025, https://opengovasia.com/2025/02/26/indonesias-ai-strategy-balancing-ai-regulation-and-innovation/
105. Indonesia Digital Transformation - International Trade Administration, accessed May 10, 2025, https://www.trade.gov/market-intelligence/indonesia-digital-transformation
106. The Latest on Southeast Asia: Indonesia joins BRICS - CSIS, accessed May 10, 2025, https://www.csis.org/blogs/latest-southeast-asia/latest-southeast-asia-indonesia-joins-brics
107. Report of the Multi-year Expert Meeting on Trade, Services and Development on its eleventh session - UNCTAD, accessed May 10, 2025, https://unctad.org/system/files/official-document/c1mem4d33_en.pdf
108. COMPUTATIONAL ANALYSIS OF THE CHINESE GOVERNMENT'S USE OF ONLINE SOCIAL NETWORKS by Zhuo Cheng A dissertation submitted to the - Niner Commons, accessed May 10, 2025, https://ninercommons.charlotte.edu/record/2633/files/Cheng_uncc_0694D_13555.pdf
109. AI in public sector: A call for digital transformation – Faculty of ..., accessed May 10, 2025, https://fia.ui.ac.id/en/ai-in-public-sector-a-call-for-digital-transformation/
110. Navigating AI and superapp adoption in Indonesia's public sector - GovInsider, accessed May 10, 2025, https://govinsider.asia/intl-en/article/navigating-ai-and-superapp-adoption-in-ind

onesias-public-sector
111. About the Journal | Indonesian Journal of International Relations - Aihii, accessed May 10, 2025, https://journal.aihii.or.id/index.php/ijir/about
112. Indonesian Journal of International Relations | ICI Journals Master List - Index Copernicus, accessed May 10, 2025, https://journals.indexcopernicus.com/search/details?id=122759&lang=en
113. MENTERI LUAR NEGERI REPUBLIK INDONESIA PERATURAN MENTERI LUAR NEGERI REPUBLIK INDONESIA NOMOR 2 TAHUN 2022 TENTANG KLASIFIKASI, accessed May 10, 2025, https://peraturan.bpk.go.id/Download/339023/Permenlu%20No.%202%20Tahun%202022.pdf