Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.6, December 2025, pp. 3000~3011
e-ISSN: 2548-6861

Sentiment Analysis of US-China Tariffs using IndoBERT and Economic Impact on Indonesia

Fitriyana Nuril Khaqqi [1]*, Ida Nurhaida [2], Lathifah Alfat [3]
[1, 2, 3] Department of Informatics, Universitas Pembangunan Jaya
[2] Center of Urban Studies, Universitas Pembangunan Jaya
fitriyana.nurilkhaqqi@student.upj.ac.id [1], ida.nurhaida@upj.ac.id [2], lathifah.alfat@upj.ac.id [3]

Article Info

Article history:

Received 2025-10-22
Revised 2025-11-21
Accepted 2025-11-26

Keyword:

IndoBERT,
Natural Language Processing,
Sentiment Analysis,
Spearman Correlation,
US-China Trade War.

ABSTRACT

The US-China trade war has influenced public perception due to its potential economic impact on developing countries like Indonesia. This study analyses Indonesian sentiment towards the tariff policies and their correlation with economic indicators. The dataset consisted of 38,739 social media comments collected through web scraping. The data were processed through data cleaning, case folding, stopword removal, normalization, and stemming. Each comment was labeled as positive, negative, and neutral. The dataset was split into 80% training and 20% testing sets, followed by an oversampling process to balance the class distribution. The data is fine-tuned using the IndoBERT model with the Python programming language. The model achieved its highest performance with an accuracy of 93.03%, precision of 93.42%, recall of 93.03%, and F1-score of 92.94%. Spearman correlation revealed a weak to moderate positive and significant correlation ($\rho = 0.434$, p-value < 0.05) between public sentiment and global soybean prices. These findings underscore the effectiveness of combining a deep learning model like IndoBERT with statistical analysis to link digital discourse to tangible economic indicators, highlighting the method's potential as a data-driven tool for policy evaluation.

I. INTRODUCTION

Economic integration encourages nations to maintain their economic independence and can potentially lead to trade conflicts [1]. Governments use import tariffs to raise prices on imported goods and protect domestic industries during a trade war [2]. The United States and China engaged in the most significant trade conflict in recent years. Both countries escalated the rivalry in 2025 by implementing new import tariff policies. These policies affected the two nations, disrupted global supply chains, and placed a disproportionate burden on developing countries involved in international trade [3].

The escalation of the US-China trade war presents a multifaceted issue for Indonesia. Indonesia can gain from trade diversion and position itself as an alternative market for either the United States or China [4]. Conversely, Indonesia faces a significant threat from an oversupply of inexpensive Chinese products that could undermine its domestic industries [5]. This twofold effect emphasizes the critical need to study public perception and to analyse the resulting domestic economic implications.

The digital age enables the monitoring of public responses and perceptions of macroeconomic issues through social media platforms. Public opinion in the digital sphere transforms into social data and influences consumer and investor sentiment [6]. Sentiment in digital spaces often reflects emotion and differs from the perspectives of economic experts. A gap between public sentiment and real economic conditions can mislead public perception and trigger public panic or misguided economic decisions [7].

The field of sentiment analysis has rapidly advanced through the use of digital data with machine learning and deep learning. A study by W. Purbaratri *et al.* [8] achieved 83% accuracy analyzing reviews of the Alpukat application using Naive Bayes, while V. Brilian Adiguna *et al.* [9] used the Random Forest algorithm and achieved 95.2% accuracy, outperforming the Support Vector Machine at 91% and Naive Bayes at 69% in film review analysis.

The research gap in previous studies lies in existing works on sentiment analysis in Indonesia being mostly focused on application reviews, entertainment content, or platform-specific consumer feedback rather than international economic policy, and no prior research integrating an Indonesian-language sentiment model with a macroeconomic-policy context. Prior studies predominantly used classical machine-learning, which tend to underperform when dealing with context-dependent sentiment because these methods rely on surface-level lexical features and cannot fully capture semantic nuances [10]. This study addresses this gap by employing IndoBERT as the primary model. IndoBERT is selected because it is pre-trained on large-scale Indonesian corpora that include news, formal articles, and socio-political narratives, which have linguistic characteristics that are closest to economic policy discussions [11]. Maharani *et al.* [12] show that IndoBERT can capture complex linguistic patterns within the domain of economic issues, while classical machine-learning models struggle to represent such features due to domain mismatch. Furthermore, studies show that IndoBERT provides higher accuracy compared to classical approaches. A study by D. Setiawan *et al.* [13] applied IndoBERT to analyse Indonesian-language comments on YouTube and achieved an accuracy of 95%, surpassing Random Forest at 85%, Support Vector Machine at 88%, and the BERT model at 92%. Another study by Fransiscus and A. S. Girsang [14] utilized IndoBERT and obtained a precision of 86%, surpassing both Support Vector Machine and Naive Bayes at 84%.

A clear understanding of public perception requires more than sentiment labels. Topic modelling methods such as BERTopic extend sentiment analysis by revealing underlying themes in textual data. A study by C. Li and X. Hu [15] has shown that applying BERTopic to datasets of scholarly literature and public opinion extracts eight primary topics and produces the top ten keywords for each topic.

Furthermore, there is a notable research gap in the absence of empirical investigation linking public sentiment to real economic indicators. This investigation is necessary because determining whether digital discourse provides meaningful market information or merely noise requires evidence of its actual economic relevance [16]. The behavioural economics framework, including the animal spirits theory [17] and the noise trader theory [18], explains how emotions and cognitive biases expressed in social media function as non fundamental information that shapes investor expectations during high uncertainty events such as trade wars [19]. This research applies statistical correlation analysis to empirically validate this theoretical relationship. A notable study examined the spearman correlation between public sentiment scores derived from Twitter and stock price fluctuations, demonstrating that positive sentiment (ρ = 0.745) and negative sentiment (ρ = 0.691) correlate statistically and significantly with stock price fluctuations [20].

This study provides a concise yet comprehensive analysis of public sentiment toward the 2025 US–China tariff escalation by enhancing linguistic precision, identifying key discourse patterns, and assessing their economic relevance. It employs IndoBERT for sentiment classification, BERTopic for thematic extraction, and a behavioral-economics lens to interpret how digital discourse aligns with economic movements. The findings offer clearer insights into how Indonesian digital communities perceive international trade tensions and support policymakers and analysts in recognising public concerns, anticipating market reactions, and formulating informed responses to global economic shocks.

II. METHOD

This study adopts a descriptive quantitative approach, structured through a systematic methodological framework.

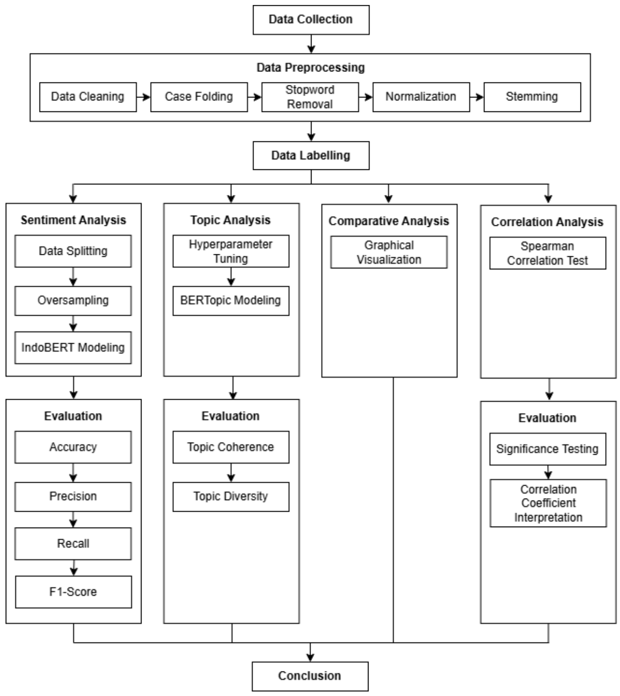

Figure 1. Flow diagram

A general overview of the research process stages is presented in Figure 1. This study conducts sentiment analysis, topic analysis, comparative analysis, and correlation analysis. The results of these analyses are evaluated to derive the key findings and obtain the conclusion.

A. Data Collection

This study uses three main components as its research data. The first component consists of 44,647 Indonesian-language public comments, collected from four major social media platforms, namely 15,286 comments from TikTok, 9,842 from Instagram, 8,315 from YouTube, and 7,204 from X. These platforms represent the largest active user bases in Indonesia [21]. Data were collected using python-based web scraping scripts, following best practices on digital trace data collection, sampling, and ethical use of publicly available content [22]. The chosen keywords, which included US China Trade War, US China Import Tariffs, and US China Tariff Hike, were selected based on a combination of trending news topics and official terminology used by international media. Content relevance was validated by filtering comments that directly responded to tariff-related news headlines, ensuring that only comments associated with tariff policy discussions remained in the dataset. The data collection period was from April 9 to May 9, 2025, a window that coincided with heightened public attention to the tariff issue. All comments were taken from the official channels of 16 national news media to ensure discussion quality and topical relevance, including Kompas, Inews, Metrotv, CNN Indonesia, CNBC Indonesia, Kumparan, Liputan6, Narasi, Tempo, Detik, Sindonews, Faktacom, Antara, Merdeka, Republika, and Okezone. The second component includes 66 expert opinions from Indonesian economists, compiled from news articles, video interviews, and social media publications. The third component is quantitative data, which includes economic indicators selected for their established connection to the impacts of US-China import tariff policies [23]. These indicators include the Jakarta Composite Index (JCI), gold prices, and global commodity prices (steel, crude oil, soybeans) from the Trading Economics website, along with stock prices of relevant commodity sector corporations—PT Krakatau Steel (KRAS), PT AKR Corporindo (AKRA), and PT Indofood Sukses Makmur (INDF)—from Yahoo Finance.

B. Data Preprocessing

The data preprocessing stage prepares the text data for model training. The procedure includes the following stages:

1) *Data Cleaning*: Data cleaning is the process of removing irrelevant elements from the text. This process involves removing URLs, punctuation, numbers, special characters, emojis, hashtags, and mentions from the text [24].

2) *Case Folding*: Case folding is the process of converting all letters in the text to lowercase to ensure word uniformity so that they are not treated as different by the model. The objective is to simplify text representation and improve the model's data processing efficiency [25].

3) Stopword Removal: Stopword removal is the process of deleting common words that do not have significant meaning in text analysis. The objective is to allow the model to focus on more meaningful words [26].

4) *Normalization*: Normalization transforms text into a consistent canonical form by correcting non-standard words, expanding abbreviations, and standardizing spelling variations. The primary objective is to ensure that the model correctly identifies lexically variant words as having the same semantic meaning [27].

5) Stemming: Stemming is the process of converting derived words into their base form by removing affixes, such as prefixes, infixes, or suffixes. This process aims to simplify analysis and reduce the data dimensionality in the model [28].

C. Data Labelling

After the pre-processing phase, the comment data underwent manual sentiment labelling. The sentiment was categorized into three primary classes, namely negative, neutral, and positive. Manual labelling was deliberately chosen to maintain the accuracy of the annotations, given that the nuances of informal language prevalent on social media are frequently challenging for automated methods to interpret correctly [29].

D. Sentiment Analysis

Sentiment analysis classified text into positive, negative, and neutral categories. This stage aimed to develop a classification model capable of recognizing sentiment within the text. The sentiment analysis process comprised several steps, including oversampling, data splitting, IndoBERT modelling, and evaluation.

1) *Oversampling*: Oversampling is a technique that resolves the issue of class imbalance. The number of instances in the minority class is augmented to ensure the model is trained on a more balanced distribution and to enhance its predictive performance [30]. The current study employs Random Oversampling (ROS), a technique that addresses class imbalance by randomly duplicating existing data points from the minority classes until the class distribution becomes balanced. The methodology employs ROS over SMOTE to preserve the linguistic authenticity of the native Indonesian comments [31].

2) *Data Splitting*: Data splitting is the process of dividing a dataset into two primary subsets, namely training data and testing data [32]. This division aims to prevent overfitting and to ensure that the model's performance can be generalized to new data. The split used a proportion of 80% for training data and 20% for testing data.

3) *IndoBERT Modelling*: The IndoBERT model is an adaptation of the BERT architecture specifically designed to understand the context of the Indonesian language. The Transformer encoder comprises a multi-head self-attention layer and a feed-forward neural network designed to capture bidirectional relationships between tokens in a sequence. The self-attention process allows the model to recognize a word's meaning based on its surrounding context, rather than solely its linear sequence. Each text token receives an embedding and a positional encoding to enable the model to understand the word's position in the sentence. Through a training process on a large Indonesian-language corpus, IndoBERT learns deep semantic representations and becomes effective for various natural language processing tasks such as sentiment analysis, text classification, and information extraction [29]. Figure 2 illustrates the Transformer encoder architecture used in IndoBERT.

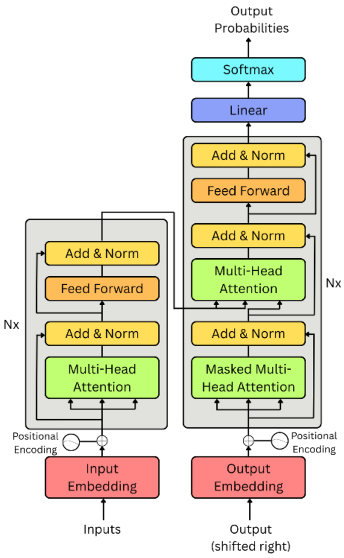

Figure 2.BERT architecture

4) *Evaluation*: The evaluation measured the model's performance to determine its effectiveness in sentiment classification [33]. This evaluation uses four primary metrics, namely:

- Accuracy represents the ratio of correct predictions to the total number of instances in the test dataset.
- Precision indicates the proportion of true positive predictions compared to all positive predictions made by the model. High precision means the model makes few false positive errors.
- Recall measures the model's ability to identify all actual positive instances within the dataset. It reflects the model's ability to capture relevant cases.
- F1-Score represents a balanced combination of precision and recall, providing a comprehensive measure of model performance.

E. *Topic Analysis*

Topic modeling identifies the main themes present in a text corpus. A hybrid approach is applied by assigning different models to each sentiment category, depending on which model performs best during the experimental phase. This approach aims to maximize the relevance and coherence of the topics extracted for each sentiment category. The topic modelling process comprises three main stages, including hyperparameter tuning, modelling with BERTopic, and evaluation.

1) *Hyperparameter Tuning*: This step tunes the hyperparameters to improve both topic coherence and topic diversity. Among the critical parameters are min_topic_size, which defines the minimum number of documents required to form a single topic, top_n_words, which determines the number of keywords that encapsulate each topic, and the embedding model [34].

2) *BERTopic Modelling*: The BERTopic model forms topic clusters based on the semantic similarity between texts. This approach helps cluster topics with similar meanings and allows dominant patterns and themes in the text data to become more clearly visible.

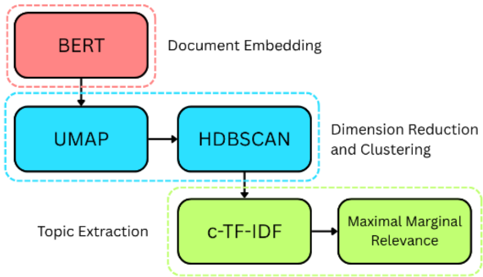

Figure 3. BERTopic architecture

Figure 3 illustrates the topic modelling process using BERTopic. This model combines transformer embeddings, dimensionality reduction (UMAP), and clustering (HDBSCAN) to generate topic clusters. Each topic then uses keywords obtained from class-based TF-IDF (c-TF-IDF) calculations to represent its content [35].

3) *Evaluation*: The evaluation assesses the model's ability to generate topics that are relevant and representative of the dataset [36]. The evaluation of the topic model uses two main metrics, namely:

- Topic Coherence measures the degree of uniformity and semantic relatedness among words within a topic. A high coherence value shows that the topic's words are semantically connected and conceptually clear.
- Topic Diversity assesses the diversity of keywords across topics to ensure that each generated topic has a unique character. A high diversity value indicates that the model does not produce overlapping topics, thereby better representing the variation of themes within the data.

F. *Comparative Analysis*

Comparative analysis is a method that compares two or more data groups, variables, or analytical results to identify similarities and differences between them [37]. This study applies comparative analysis to contrast public sentiment with the opinion of economic experts. Both variables are visualized in graphs to clearly and informatively illustrate the differences and trends between them.

G. *Behavioural Economics*

Figure 4 illustrates the behavioural economics framework designed to operationalize the theoretical link between public sentiment and economic indicators. The framework utilizes unstructured public opinion from social media as the primary input for analysis. The animal spirits theory [17] guides the sentiment analysis stage to quantify abstract psychological factors like collective optimism or fear.

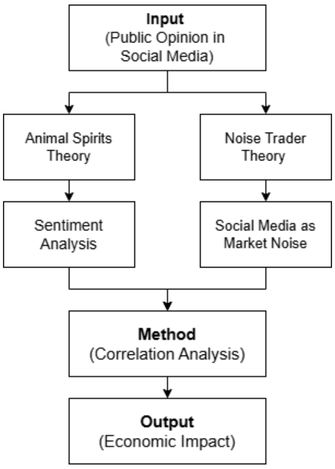

Figure 4. Behavioural Economics Framework.

Concurrently, the noise trader theory [18] is applied to posit that while social media discourse represents market noise distinct from fundamental data, this aggregated noise significantly influences perceived value and behaviour. Consequently, public sentiment drives macroeconomic impact by shaping the collective consumption and investment decisions that fundamentally determine aggregate demand [38]. Specifically, collective optimism measured via social media causally boosts household consumption, thereby supporting macroeconomic growth [39]. The final correlation analysis examines the relationship between these quantified sentiments and economic variables to determine the resultant economic impact.

H. Correlation Analysis

Correlation analysis measures the correlation between public sentiment scores and various economic indicators on a daily basis over a one-month period. This stage aims to determine the extent to which changes in public sentiment correlate with ongoing economic dynamics.

Before conducting the Spearman correlation, the daily public sentiment score is calculated from the distribution of sentiment categories. Each comment receives a score based on its classification, and the daily sentiment score is computed using the following formula:

$$\text{Score} = \frac{(\text{negative} \times -1 + \text{neutral} \times 0 + \text{positive} \times 1)}{\text{total comments}} \quad (1)$$

The correlation analysis process involves two main steps, namely:

1) *Spearman Correlation*: Spearman correlation is a non-parametric statistical method that measures the strength and direction of the correlation between two variables based on ranked data [40]. Public sentiment scores and daily commodity prices typically exhibit non-normal distributions and outliers [41]. The strategic selection of this method aims to isolate fundamental trends from the noise of market volatility that often masks complex macroeconomic dynamics [42]. This rank-based approach provides superior statistical validity over parametric methods that remain highly sensitive and prone to bias from extreme outliers inherent in these volatile datasets [43]. The technique ensures robust estimation of the relationship strength and prevents the overestimation of correlation significance caused by non-normally distributed transient data spikes.

2) *Evaluation*: The evaluation of correlation results assesses the significance and strength of the correlation between two variables tested using Spearman correlation. This stage involves two main aspects, namely:

- A significance test ensures that the correlation found between two variables is statistically significant and not due to chance. A significance value (p-value) less than 0.05 indicates that the resulting correlation is significant [44].
- Interpretation of the correlation coefficient (ρ) determines the direction and strength of the correlation between variables. A coefficient value close to 1 indicates a strong positive correlation, a value close to -1 indicates a strong negative correlation, while a value near 0 indicates a weak or insignificant correlation [45].

III. Result and Discussion

A. Dataset

TABLE 1
RAW PUBLIC COMMENT DATA

Platform	Date	Comment
Tiktok	April 9, 2025	*mau dikenakan 200% pun. nga masalah. barang China tetap laku di Amerika. sedangkan barang Amerika kalau kena 100persen aja. nga ada yg mau beli.* 😂 Even if a 200% tariff is imposed, it doesn't matter. Chinese goods will still sell well in America. Meanwhile, if American goods get even a 100% tariff, no one will want to buy them. 😂
Youtube	April 15, 2025	*Rakyat USA sbg Pembeli juga barang IMPORT akan Berat dan membeli mahal..kayaknya akan ada demo beasar semua negara Bagian USA* 😅 The people of the USA, as buyers of imported goods, will also be burdened and have to buy at high prices... it seems there will be massive protests across all U.S. states 😅
Instagram	April 17, 2025	*Jgn sampe perang ekonomi berubah jd perang senjata krn ego trump* 😂 Hopefully the economic war doesn't turn into an armed war because of Trump's ego 😂
X	April 25, 2025	*ini baru negara besar dan berdikari, bukan omon omon* Now this is what you call a great and self-reliant nation, not just empty talk

The dataset comprises three main components, namely public comment data, economic expert opinions, and economic indicators. The first component is public comment data, serving as the primary source for sentiment analysis. Table 1 presents several representative examples of the raw public comment data.

The second component is economic expert opinion data. Table 2 presents representative examples of the raw economic expert opinion data used for comparison with public sentiment.

TABLE II
RAW ECONOMIC EXPERT OPINION DATA

Source	Economic Expert	Opinion
Website valoranews	Direktur Eksekutif CORE Indonesia, Mohammad Faisal	*Indonesia menghadapi risiko serius akibat perang dagang AS-Tiongkok.* Indonesia faces serious risks due to the US-China trade war.

The third component is economic indicator data. Table 3 presents the economic indicator data.

TABLE III
ECONOMIC INDICATOR DATA

Economic Indicator	Value			
	April 09, 2025	**April 10, 2025**	**…**	**May 09, 2025**
Steel Price (USD/Ton)	920	930	…	890
Crude Oil Price (USD/Bbl)	62.35	60.07	…	61.02
Soybean Price (USD/Bu)	1,012.8	1029	…	1,051.8
Gold Price (IDR)	1,670, 200	1,727, 518	…	1,764,52 4
Stock Price of PT Krakatau Steel (KRAS)	106	111	…	126
Stock Price of PT AKR Corporindo (AKRA)	900	1,100	…	1,245
Stock Price of PT Indofood Sukses Makmur (INDF)	6,675	7,075	…	7,975
Jakarta Composite Index (JCI)	5,967.99	6,254.02	…	6,832.8

B. Data Preprocessing

The first stage is data cleaning, which begins with selecting the most relevant columns for analysis, specifically those containing comment text. Other columns are removed as they do not directly affect sentiment analysis. Additionally, missing values and duplicate data are eliminated to ensure the accuracy and consistency of the data used. After filtering, the text is cleaned of unnecessary elements using the Regular Expression (regex) library in Python. Subsequently, case folding converts all characters to lowercase. This operation is performed automatically using the .lower() function in Python. The next stage is stopword removal. This process is conducted using the Indonesian stopword list from the NLTK library in Python. The subsequent phase is normalization, which aims to transform non-standard words and abbreviations into their standard forms. This normalization is conducted using a manually curated dictionary for social media slang, assisted by the replace function from the pandas library. The last phase is stemming, a process that reduces words with affixes to their fundamental form. This stemming operation is conducted using the Sastrawi library in Python. Following the preprocessing phase, the dataset was reduced to 38,739 valid comments. Table 4 presents the comment text data resulting from the data preprocessing stage.

TABLE IV
RESULT OF DATA CLEANING

Stage	Result
Original Text	*Perang dagang kali ini Amerika kalah dan dimenangkan oleh China.* This trade war, the United States lost and China won.
Data Cleaning	*Perang dagang kali ini Amerika kalah dan dimenangkan oleh China* This trade war the United States lost and China won
Case Folding	*perang dagang kali ini amerika kalah dan dimenangkan oleh china* this trade war the united states lost and china won
Stopword Removal	*perang dagang amerika kalah dimenangkan china* trade war united states lost won china
Normalization	*perang dagang amerika kalah menang china* trade war united states lost won china
Stemming	*perang dagang amerika kalah menang china* trade war united state lose win china

C. Data Labelling

The data labelling process assigns a sentiment label to each comment text. Comments with a supportive or optimistic tone were categorized as positive, comments that were critical or pessimistic were categorized as negative, while those that were informative, neutral, or ambiguous were placed in the neutral category. The three authors independently labeled sentiments based on predefined guidelines to mitigate subjectivity. The study quantified inter-annotator agreement using Fleiss' Kappa to ensure the reliability and consistency of the manual labeling process [46]. This metric yielded a score of 0.92, which indicated strong agreement. Table 5 displays examples of the comment text data after the data labelling stage.

TABLE V
RESULT OF DATA LABELLING

Comment	Label
amerika serikat kewalahan perang dagang china naik tarif peras ekonomi amerika serikat jadi preman the united states is overwhelmed by the trade war with china raising tariffs squeezing the u.s. economy and turning into a thug	Negative
cina santai biar mau naik sangat yakin cina sangat kuat negara sedang maju segala bidang china is relaxed even if it rises very confident china is a very strong developing country in all fields	Positive
dampak tingkat kurs mata uang the impact of the exchange rate level	Neutral

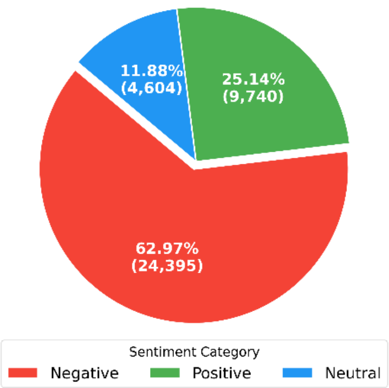

Figure 5. Pie chart of sentiment distribution

Figure 5 presents a pie chart that visually represents the distribution of sentiment in the comment data. The sentiment distribution results show that most comments fall into the negative category with a total of 24,395 data points.

D. *Implementation of the IndoBERT Model*

The initial dataset analysis reveals a significant imbalance between the sentiment classes. Random Oversampling (ROS) was applied to address the significant class imbalance. This technique duplicates instances from the minority sentiment classes (positive and neutral) at random until their frequency matches the majority class (negative). The implementation uses the RandomOverSampler module from the imbalanced-learn (imblearn) Python library.

The oversampling procedure is applied after the data-splitting stage, ensuring that only the training dataset is affected and preventing data leakage. The train_test_split function from the scikit-learn library partitioned the original dataset into 80% for training and 20% for testing. The testing set remained purely original and unseen, ensuring an unbiased evaluation of the IndoBERT model.

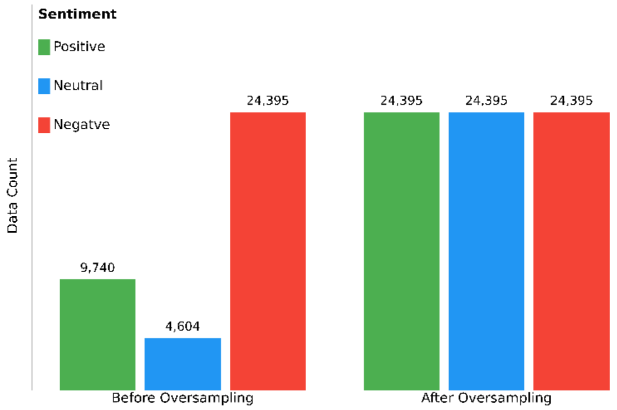

Figure 6. Chart illustrating the change in sentiment data distribution before and after the oversampling process.

Figure 6 presents a comparison of the data count for each sentiment class before and after the oversampling process. This process increases the amount of data for the positive and neutral classes from 9,740 and 4,604 to 24,395, matching the amount of data in the negative class.

Subsequently, the IndoBERT model applies the training parameter configuration detailed in Table 6.

TABLE VI
INDOBERT MODEL TRAINING PARAMETERS

Parameter	Value
Base Model	IndoBERT-base-p1
Learning Rate	2e-5
Batch Size	16
Maximum Sequence Length	512
Number of Epochs	5
Optimizer	AdamW
Loss Function	Cross-Entropy Loss

After the model configuration was established, three experiments compared the effect of the data balancing technique on sentiment classification performance as follows:

- Experiment 1: Model training uses the pre-processed data without any balancing technique.
- Experiment 2: Model training uses the pre-processed data, incorporating class weighting into the loss function.
- Experiment 3: Model training uses the oversampled data.

The evaluation of the model's performance in all three experiments uses accuracy, precision, recall, and F1-score metrics. The results of the performance comparison for the three experiments appear in Table 7.

TABLE VII
RESULT OF INDOBERT FINE-TUNING EVALUATION

Experiment	Accuracy	Precision	Recall	F1-Score
Experiment 1	0.8242	0.8223	0.8242	0.8223
Experiment 2	0.8726	0.8730	0.8726	0.8723
Experiment 3	0.9303	0.9342	0.9303	0.9294

The model utilizing the oversampling technique demonstrates strong performance across all evaluation metrics. The enhancement in these four metrics signifies that oversampling effectively enhances the model's ability to learn from a more balanced dataset. This finding also substantiates that a data-level intervention that enriches the sample size of minority classes constitutes a more effective and robust approach than an algorithm-level intervention such as class weighting.

The comparative approach also conducted a straightforward baseline comparison against Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Naive Bayes. The evaluation process employed these models to evaluate the transformer model's efficacy against conventional and deep learning approaches. Naive Bayes configured using TF-IDF representation with a smoothing parameter α = 1.0. The SVM baseline utilizing a linear kernel and a regularization value of C = 1.0. Meanwhile, the LSTM model employed a single LSTM layer with 128 units and a 300-dimensional embedding. The results are presented in Table 8. The IndoBERT model consistently demonstrates the highest performance across all evaluation metrics. The IndoBERT model is thus the optimal choice when the primary goal is achieving a high overall classification score.

TABLE VIII
PERFORMANCE COMPARISON OF INDOBERT AGAINST BASELINE MODEL

Model	Accuracy	Precision	Recall	F1-Score
Naive Bayes	0.732	0.694	0.758	0.724
SVM	0.823	0.842	0.787	0.814
LSTM	0.872	0.893	0.854	0.873
IndoBERT (Experiment 3)	0.930	0.934	0.930	0.929

E. *Implementation of the BERTopic Model*

The subsequent stage implements the BERTopic model to identify the primary topics arising from public comments. During this process, hyperparameter tuning determines the optimal parameter combination for each sentiment category. The paraphrase-multilingual-MiniLM-L12-v2 model, with a min_topic_size of 15, yields reasonably good coherence and diversity values for both positive and negative topics. For the positive category, it achieves a coherence score of 0.527 and a diversity score of 0.815, whereas for the negative category, the coherence and diversity scores are 0.412 and 0.777, respectively. The indobenchmark/indobert-base-p2 model demonstrates the strongest performance for the neutral category, with a topic coherence of 0.643 and a topic diversity of 0.900. The third model is used for topic modelling of the neutral sentiment, while the second model is applied to the positive and negative sentiments. The complete results of the experiments for various embedding and parameter combinations appear in Table 9.

TABLE IX
RESULT OF BERTOPIC MODEL IMPLEMENTATION

Model and Parameter	Sentiment	Topic Coherence	Topic Diversity
1. min_topic_size = 10 2. nr_topics = "auto" 3. top_n_words = 10 4. model = Paraphrase -multilingual-MiniLM -L12-v2	Positive	0.373	0.908
	Negative	0.347	0.911
	Neutral	0.367	0.902
1. min_topic_size = 15 2. nr_topics = "auto" 3. top_n_words = 15 4. model = Paraphrase -multilingual-MiniLM -L12-v2	Positive	0.527	0.815
	Negative	0.412	0.777
	Neutral	0.453	0.765
1. min_topic_size = 10 2. nr_topics = "auto" 3. top_n_words = 10 4. model = Indobench mark/indobert-base-p2	Positive	0.442	0.756
	Negative	0.408	0.738
	Neutral	0.643	0.900
1. min_topic_size = 15 2. nr_topics = "auto" 3. top_n_words = 15 4. model = Indobench mark/indobert-base-p2	Positive	0.479	0.772
	Negative	0.425	0.709
	Neutral	0.432	0.756

Based on the topic modelling results for the positive sentiment category, three dominant topics were identified, namely the dynamics of economic power between China and America, Indonesia's hope and courage, and support for China or America. Figure 7 illustrates the top keywords that construct and characterize each of these topics.

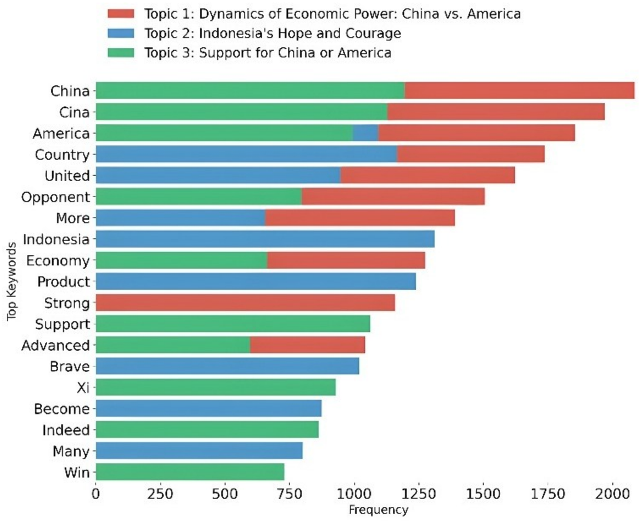

Figure 7. Result of positive topic modelling.

The topic modelling results for the negative sentiment category reveal three major topics, namely the impact of the US-China trade conflict, concerns of escalation into a world war, and criticism of President Donald Trump. Figure 8 displays the top keywords that form and describe each of these topics.

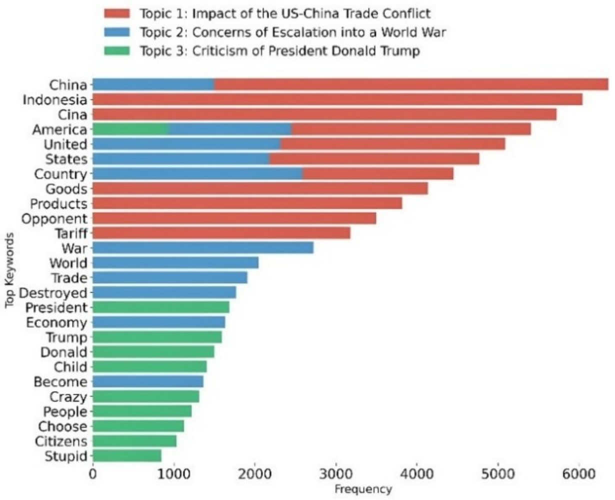

Figure 8. Result of negative topic modelling.

The topic modelling results for the neutral sentiment category indicate two primary topics, namely general discussion of the US-China trade conflict and chronology of retaliatory import tariffs. Figure 9 illustrates the top keywords that form and represent each of these topics.

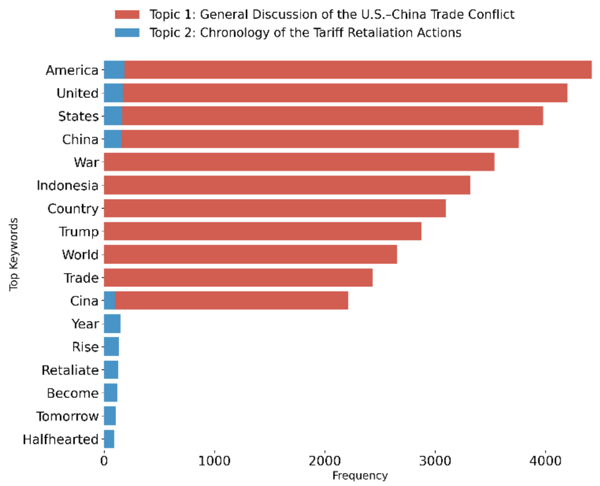

Figure 9. Result of neutral topic modelling.

F. *Results of Comparing Public Sentiment with Economic Experts*

Public commentary is predominantly negative (62.97%), followed by positive (25.14%) and neutral (11.88%) responses. Meanwhile, economic experts express 54.55% negative, 12.12% positive, and 33.33% neutral sentiment. This comparison indicates that both the public and economic experts generally respond negatively to the increase in US import tariffs on China. The prevalence of negative sentiment in both cohorts highlights that the policy is perceived as potentially leading to adverse effects or raising concerns from both lay and expert perspectives, while a key difference is the distribution of neutral sentiment, which is proportionally higher among experts. This indicates that experts approach the issue with a more analytical and measured perspective. In contrast, public sentiment tends to be more reactive and emotional, focusing on the negative "war" narrative that is perceived as harmful, which results in a higher proportion of negative sentiment. Figure 10 presents a visual comparison of the sentiment analysis results between the public and economic experts.

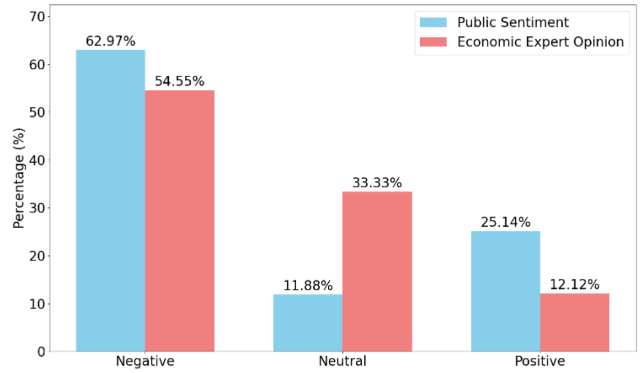

Figure 10. Graph comparing public sentiment with the views of economic experts.

G. *Spearman Correlation Analysis between Public Sentiment and Economic Indicators*

The Spearman correlation coefficient examines the relationship between daily public sentiment and various economic variables. Consequently, each day is represented by a single value that reflects the prevailing public sentiment. A portion of these daily scores is presented in Table 10.

TABLE X
DAILY PUBLIC SENTIMENT SCORE

Sentiment Score	Date
-0,394	April 09, 2025
-0,353	April 10, 2025
…	…
-0,3795	May 09, 2025

Python is used to conduct the Spearman correlation analysis, employing the pandas library and the scipy.stats module to compute the correlation coefficient and its corresponding p-value. The test results indicate that only global soybean prices have a p-value of 0.049, which makes it statistically significant. The other variables have a p-value $\geq$ 0.05, so their association with public sentiment is not statistically significant. Table 11 presents the resulting Spearman correlation coefficients and p-values for the correlation between public sentiment scores and various economic indicators.

TABLE XI
RESULT OF SPEARMAN CORRELATION AND P-VALUE

Economic Indicator	Spearman (ρ)	p-value
Steel Price	0.122	0.597
Crude Oil Price	0.032	0.889
Soybean Price	0.434	0.049
Gold Price	-0.066	0.775
Stock Price of PT Krakatau Steel (KRAS)	-0.203	0.378
Stock Price of PT AKR Corporindo (AKRA)	-0.175	0.448
Stock Price of PT Indofood Sukses Makmur (INDF)	-0.010	0.964
Jakarta Composite Index (JCI)	-0.194	0.399

The subsequent step is to interpret the correlation coefficient (ρ) to gauge the strength of the association between the variables after establishing statistical significance. The interpretation is as follows:

- The global soybean price exhibits a weak to moderate positive association with public sentiment. The result indicates that a decline in public sentiment tends to coincide with a decrease in global soybean prices.
- Variables demonstrating a weak association include the KRAS stock price, the JCI, the AKRA stock price, and global steel prices. These findings suggest that, although minor trends exist, none of these associations achieves statistical significance.
- Variables categorized as having a very weak or negligible association comprise gold prices, crude oil prices, and the INDF stock price. The small coefficient values and high p-values indicate an absence of meaningful association, implying that any observed fluctuations are likely attributable to random variation rather than systematic patterns.

The positive correlation found between public sentiment and global soybean prices implies that both variables move in tandem as a deterioration in sentiment coincides with a decline in prices. This synchronization stems from the historical status of soybeans as the primary retaliatory commodity in US-China trade tensions [47]. Media framing ensures that negative trade news triggers immediate public pessimism while simultaneously signaling reduced export demand to market participants [48]. These parallel reactions cause sentiment scores and commodity prices to decline in tandem during periods of high volatility. Noise trader theory supports this interpretation because collective pessimism on social media reflects the bearish psychology of the market [18]. External confounding variables such as weather, harvest yields, and logistics undoubtedly affect physical supply levels independently of social discourse [49]. The persistence of the sentiment-price correlation despite these factors indicates that social mood specifically captures the distinct volatility caused by trade policy shocks rather than physical constraints.

The absence of significant correlations in other indicators can be explained by the influence of broader global economic factors during the observation period. Short term political events like tariff announcements often have limited impact on diversified market indices because these indices are also affected by global market sentiment, monetary policy expectations, and sector specific conditions [50].

IV. CONCLUSION

This study found that the IndoBERT model was effective in analysing public sentiment on social media regarding the US-China tariff policies. The model achieves its highest performance with an accuracy of 93.03%, a precision of 93.42%, a recall of 93.03%, and an F1-score of 92.94%, demonstrating its robustness in sentiment classification. The analysis shows that public sentiment remains predominantly negative, which corresponds with economic experts. BERTopic clusters the social media comments and identifies that positive comments focus on China's economic strength and opportunities for Indonesia, negative comments highlight global economic impact and national economic risk, and neutral comments address policy chronology and trade relations development. The correlation analysis shows that the global soybean price correlates significantly with public sentiment, exhibiting a weak-to-moderate positive correlation.

This study acknowledges several limitations regarding dataset bias, variable scope, and methodological approach. The dataset only includes social media comments, so it mainly reflects the views of active digital users rather than the broader Indonesian population or institutional investors. The one-month Spearman correlation observation period is too short to capture the full impact of public sentiment on macroeconomic indicators. The economic variables are limited to certain commodity and stock prices, leaving out wider indicators like inflation, trade balance, GDP growth, global economic conditions, interest rate expectations, or supply shocks. The Spearman correlation is therefore only an exploratory measure of association, not proof of causation. Future research should expand the scope of macroeconomic variables and extend the observation period to produce more reliable findings. In addition, employing multivariate econometric models such as vector autoregression, autoregressive distributed lag, or regression with control variables would allow for a more robust identification of the drivers of economic changes and help isolate the specific influence of public sentiment from other macroeconomic shocks.

REFERENCES

[1] Q. Gallea and D. Rohner, "Globalization mitigates the risk of conflict caused by strategic territory," *Proc Natl Acad Sci U S A*, vol. 118, no. 39, 2021, doi: 10.1073/pnas.2105624118.

[2] D. Furceri, S. A. Hannan, J. D. Ostry, and A. K. Rose, "Are tariffs bad for growth? Yes, say five decades of data from 150 countries," *J Policy Model*, vol. 42, no. 4, 2020, doi: 10.1016/j.jpolmod.2020.03.009.

[3] C. Nurul Fadilah, E. Jumhana, T. Salsabilla, N. Septiani, and R. Zamir, "Perang Dagang Dan Proteksionisme: Kajian Ekonomi

Politik Atas Kebijakan Tarif Dalam Hubungan Dagang Bilateral," *Jurnal Riset Ilmu Hukum*, vol. 2, pp. 220–229, 2025, doi: 10.62383/mahkamah.v2i3.860.

[4] H. Y. Jong, "The Art of Trade War: Spurring Investments in Indonesia Amidst the US–China Trade War," *Global Journal of Emerging Market Economies*, vol. 14, no. 2, 2022, doi: 10.1177/09749101211034110.

[5] R. Purwono, U. Heriqbaldi, M. A. Esquivias, and M. K. Mubin, "The American–China Trade War and Spillover Effects on Value-Added Exports from Indonesia," *Sustainability (Switzerland)*, vol. 14, no. 5, 2022, doi: 10.3390/su14053093.

[6] P. N. Reis, A. P. S. Pinto, and A. Guimarães, "'The influence of consumer, manager, and investor sentiment on US stock market returns,'" *Investment Management and Financial Innovations*, vol. 22, no. 1, pp. 231–256, 2025, doi: 10.21511/imfi.22(1).2025.18.

[7] R. Agrawal, S. Mishra, and P. Mathur, "Psychological Impact Of Social Media On Investor Confidence," *South Eastern European Journal of Public Health (SEEJPH)*, vol. XXV, pp. 5–12, 2024.

[8] W. Purbaratri, H. D. Purnomo, D. Manongga, I. Setyawan, and H. Hendry, "Sentiment Analysis of e-Government Service Using the Naive Bayes Algorithm," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 23, no. 2, pp. 441–452, Mar. 2024, doi: 10.30812/matrik.v23i2.3272.

[9] V. Brilian Adiguna, M. Aqqad, J. S. Pt, and H. R. Dp, "Improving Performance Sentiment Analysis Movie Review Film using Random Forest with Feature Selection Information Gain," *International Journal of Artificial Intelegence Research*, vol. 8, no. 1, p. 1, 2024, [Online]. Available: http://ijair.id

[10] A. Karanikola, G. Davrazos, C. M. Liapis, and S. Kotsiantis, "Financial sentiment analysis: Classic methods vs. deep learning models," *Intelligent Decision Technologies*, vol. 17, no. 4, 2023, doi: 10.3233/IDT-230478.

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

[12] N. P. I. Maharani, A. Purwarianti, Y. Yustiawan, and F. C. Rochim, "Domain-Specific Language Model Post-Training for Indonesian Financial NLP," in *Proceedings of the International Conference on Electrical Engineering and Informatics*, 2023. doi: 10.1109/ICEEI59426.2023.10346625.

[13] D. Setiawan, D. Utari Iswavigra, and E. Anggiratih, "Implementation of IndoBERT for Sentiment Analysis of the Constitutional Court's Decision Regarding the Minimum Age of Vice Presidential Candidates," *Scientific Journal of Informatics*, vol. 12, no. 3, 2025, doi: 10.15294/sji.v12i3.26360.

[14] Fransiscus and A. S. Girsang, "Sentiment Analysis of COVID-19 Public Activity Restriction (PPKM) Impact using BERT Method," *International Journal of Engineering Trends and Technology*, vol. 70, no. 12, 2022, doi: 10.14445/22315381/IJETT-V70I12P226.

[15] C. Li and X. Hu, "Medical Artificial Intelligence in Scholarly and Public Perspective: BERTopic-Based Analysis of Topic-Sentiment Collaborative Mining," *Data Science and Informetrics*, May 2025, doi: 10.1016/J.DSIM.2025.05.001.

[16] I. Fitriani and A. Amalyah Agus, "Tweeting the economy: Analyzing social media sentiments and macroeconomic indicators," *Journal of Innovation in Business and Economics*, vol. 7, no. 02, 2023, doi: 10.22219/jibe.v7i02.28227.

[17] G. La Cava, "Smells like animal spirits: the sensitivity of corporate investment to sentiment," *Macroecon Dyn*, vol. 27, no. 8, 2023, doi: 10.1017/S1365100522000608.

[18] J. Yoon and G. Oh, "Investor herding behavior in social media sentiment," *Front Phys*, vol. 10, 2022, doi: 10.3389/fphy.2022.1023071.

[19] S. Daida, Y. Yadav, K. Kaustubh Sontakke, A. Bagkar, K. Arvind Sontakke, and A. Pathak, "Behavioral Finance and Investor Psychology: Understanding Market Volatility in Crisis Scenarios," *Advances in Consumer Research*, vol. 2, no. 5, pp. 851–863, Oct. 2025, [Online]. Available: https://acr-journal.com/

[20] M. N. Ibnu Sina and E. B. Setiawan, "Stock Price Correlation Analysis with Twitter Sentiment Analysis Using The CNN-LSTM Method," *sinkron*, vol. 8, no. 4, pp. 2190–2202, Oct. 2023, doi: 10.33395/SINKRON.V8I4.12855.

[21] A. P. Nugraha, A. Kurnia, P. Irawan, P. Putra, A. Rahman, and D. Dikrurahman, "The Impact of Social Media on Social Interaction and Self-Identity in Indonesian Society," *Journal of Social Research*, vol. 3, no. 9, Aug. 2024, doi: 10.55324/JOSR.V3I9.2254.

[22] J. Ohme *et al.*, "Digital Trace Data Collection for Social Media Effects Research: APIs, Data Donation, and (Screen) Tracking," *Commun Methods Meas*, vol. 18, no. 2, 2024, doi: 10.1080/19312458.2023.2181319.

[23] L. N. Sulistiyowati and S. V. Pratama, "Indonesia's macroeconomic conditions during United States-China trade war," *Jurnal Ekonomi dan Bisnis*, vol. 26, no. 2, pp. 509–530, Dec. 2023, doi: 10.24914/JEB.V26I2.6252.

[24] W. Ben Khalifa, "Using Natural Language Processing Techniques for Indexing and Analyzing Archive Documents," *Journal of Posthumanism*, vol. 5, no. 1, pp. 1520–1540, May 2025, doi: 10.63332/joph.v5i1.1394.

[25] M. Siino, I. Tinnirello, and M. La Cascia, "Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers," *Inf Syst*, vol. 121, pp. 306–4379, 2024, doi: 10.1016/j.is.2023.102342.

[26] J. Kaur and P. K. Buttar, "A Systematic Review on Stopword Removal Algorithms," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 4, no. 4, 2018.

[27] A. R. Lubis and M. K. M. Nasution, "Twitter Data Analysis And Text Normalization In Collecting Standard Word," *Journal of Applied Engineering and Technological Science*, vol. 4, no. 2, 2023, doi: 10.37385/jaets.v4i2.1991.

[28] M. D. Purbolaksono, F. D. Reskyadita, Adiwijaya, A. A. Suryani, and A. F. Huda, "Indonesian text classification using back propagation and sastrawi stemming analysis with information gain for selection feature," *Int J Adv Sci Eng Inf Technol*, no. 1, 2020, doi: 10.18517/ijaseit.10.1.8858.

[29] A. H. Nasution and A. Onan, "ChatGPT Label: Comparing the Quality of Human-Generated and LLM-Generated Annotations in Low-Resource Language NLP Tasks," *IEEE Access*, vol. 12, pp. 71876–71900, 2024, doi: 10.1109/ACCESS.2024.3402809.

[30] L. A. Martini, G. A. Pradipta, and R. R. Huizen, "Analysis of the Impact of Data Oversampling on the Support Vector Machine Method for Stroke Disease Classification," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 2, pp. 404–421, Mar. 2025, doi: 10.35882/JEEEMI.V7I2.698.

[31] G. Eom and H. Byeon, "Searching for Optimal Oversampling to Process Imbalanced Data: Generative Adversarial Networks and Synthetic Minority Over-Sampling Technique," *Mathematics*, vol. 11, no. 16, 2023, doi: 10.3390/math11163605.

[32] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," *Jurnal Sains dan Informatika*, vol. 9, no. 1, pp. 19–28, Jun. 2023, doi: 10.34128/JSI.V9I1.622.

[33] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci Rep*, vol. 14, no. 1, p. 6086, Dec. 2024, doi: 10.1038/S41598-024-56706-X.

[34] S. Syahrial and R. P. F. Afidh, "Fine-Tuning Topic Modelling: A Coherence-Focused Analysis of Correlated Topic Models," *Infolitika Journal of Data Science*, vol. 2, no. 2, pp. 82–87, Nov. 2024, doi: 10.60084/ijds.v2i2.236.

[35] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," Mar. 2022, [Online]. Available: http://arxiv.org/abs/2203.05794

[36] A. B. Dieng, F. J. R. Ruiz, and D. M. Blei, "Topic Modeling in Embedding Spaces", doi: 10.1162/tacl.

[37] S. Rivas, Y. Hernandez, R. Urraca, and P. Barbosa, "A comparative analysis to depict underlying attributes that might determine successful implementation of local adaptation plans," *Environ Sci Policy*, vol. 117, pp. 25–33, Mar. 2021, doi: 10.1016/J.ENVSCI.2020.12.002.

[38] I. Fitriani and A. A. Agus, "Tweeting the economy: Analyzing social media sentiments and macroeconomic indicators," *Journal of Innovation in Business and Economics*, vol. 7, no. 02, pp. 129–138, Nov. 2023, doi: 10.22219/JIBE.V7I02.28227.

[39] C. A. Makridis, "The social transmission of economic sentiment on consumption," *Eur Econ Rev*, vol. 148, p. 104232, Sep. 2022, doi: 10.1016/J.EUROECOREV.2022.104232.

[40] S. K. Eden, C. Li, and B. E. Shepherd, "Spearman-like correlation measure adjusting for covariates in bivariate survival data," *Biom J*, vol. 65, no. 8, p. e2200137, Dec. 2023, doi: 10.1002/BIMJ.202200137.

[41] M. Winistörfer and I. Zhdankin, "Measure of Dependence for Financial Time-Series," *SSRN Electronic Journal*, Nov. 2023, doi: 10.2139/SSRN.4641504.

[42] L. J. Yacoubian, G. S. Ketenchian, G. H. F. de Carvalho, and K. Merdinian, "Market Dynamics in Diverging Economies: A Comparative Study of the MERVAL and S&P 500 Indices," *Revista JRG de Estudos Acadêmicos*, vol. 5, no. 11, pp. 623–632, Dec. 2022, doi: 10.55892/JRG.V5I11.2364.

[43] A. F. Suherman, P. P. Lisnaeni, S. A. Izqiatullailiyah, T. Herlinawati, and A. Ahman, "A Comparative Analysis Of Spearman And Pearson Correlation Using SPSS," *OPTIMA: Journal Of Guidance and Counseling*, vol. 5, no. 1, pp. 1–15, Mar. 2025, Accessed: Nov. 21, 2025. [Online]. Available: https://ejournal.upi.edu/index.php/optima/article/view/79415

[44] G. W. Imbens, "Statistical significance, p-values, and the reporting of uncertainty," 2021. doi: 10.1257/jep.35.3.157.

[45] R. J. Janse *et al.*, "Conducting correlation analysis: important limitations and pitfalls," *Clin Kidney J*, vol. 14, no. 11, p. 2332, Nov. 2021, doi: 10.1093/CKJ/SFAB085.

[46] F. Moons and E. Vandervieren, "Measuring agreement among several raters classifying subjects into one or more (hierarchical) categories: A generalization of Fleiss' kappa," *Behav Res Methods*, vol. 57, no. 10, Oct. 2025, doi: 10.3758/s13428-025-02746-8.

[47] A. Bandyopadhyay and P. Rajib, "The impact of Sino–US trade war on price discovery of soybean: A double-edged sword?," *Journal of Futures Markets*, vol. 43, no. 7, 2023, doi: 10.1002/fut.22415.

[48] S. Qu, L. Rodriguez, and E. Cork, "Media frames and farmers' support for the 2018 U.S. trade dispute with China," *Journal of International Communication*, vol. 30, no. 1, 2024, doi: 10.1080/13216597.2024.2311714.

[49] F. D. da C. Mello, P. Kumar, and E. G. Sperandio Nascimento, "Advancements in Soybean Price Forecasting: Impact of AI and Critical Research Gaps in Global Markets," *MDPI*, vol. 12, no. 11, Nov. 2024, doi: 10.3390/economies12110310.

[50] M. Sajid Nazir, H. Younus, A. Kaleem, and Z. Anwar, "Impact of political events on stock market returns: empirical evidence from Pakistan," *Journal of Economic and Administrative Sciences*, vol. 30, no. 1, 2014, doi: 10.1108/JEAS-03-2013-0011.