Sentiment Analysis with Valency and Tendency Functions Using Bert Sequence Model

THESIS

Submitted as fulfilment of the requirements for the completion of

Master of Computer Science Program

LUTFIL KHAKIM

20210130035

MASTER OF COMPUTER SCIENCE PROGRAM

SCHOOL OF COMPUTER SCIENCE

2024

STATEMENT OF AUTHENTICITY

The undersigned below:

Name	:	Lutfil Khakim
ID of student	:	20210130035
Faculty	:	Computer Science
The Tittle of Thesis	:	Sentiment Analysis with Valency and Tendency Functions using BERT Sequence Model

Stating truthfully that this thesis has nothing in common with other thesis. Thus this statement is made without coercion from any party. If this statement is not true, it will be sanctioned by the faculty leader.

Sukabumi, August 2024

Writer,

Lutfil Khakim

NIM. 20190130035

APPROVAL OF THESIS

Title	:	Sentiment Analysis with Valency and Tendency Functions using BERT Sequence Model
Name	:	Lutfil Khakim
ID of student Faculty	:	20210130035

This thesis has been reviewed and approved

Sukabumi, August 2024

Head of Study Program,	Supervisor,
Prof. Ir. Teddy Mantoro, M.Sc., PhD	Prof. Ir. Media Anugerah Ayu, M.Sc., PhD
NIDN. 0323096491	NIDN. 0315046903

THESIS APPROVAL

Title	:	Sentiment Analysis with Valency and Tendency Functions using BERT Sequence Model
Name	:	Lutfil Khakim
ID of student Faculty	:	20210130035

This Thesis has been tested and defended in front of the Board of Examiners in Thesis session on 18 August 2024. In our review, this Thesis is adequate in terms of quality for the purpose of awarding the Master of Computer Degree.

Sukabumi, August 2024

Supervisor 1,	Examiner 1,
Prof. Ir. Teddy Mantoro, M.Sc., PhD	Deshinta Arrova Dewi, M.Sc., PhD.
NIDN. 0323096491	NIDN. 0416127102
Supervisor 2,	Examiner 2,
Prof. Ir. Media Anugerah Ayu, M.Sc., PhD	Rahmadya T. Handayanto, Ph.D.
NIDN. 0315046903	NIDN. 0307067601

PUBLICATION APPROVAL

As a member of the academic community of Nusa Putra University, i undersigned:

Name	:	Lutfil Khakim
ID of student	:	20210130035
Study Program	:	Computer Science
Type of Work	:	Thesis

For the sake of scientific development, agree to grant to the University of Nusa Putra the Non-Exclusive Royality-Free Right for my scientific work entitled: **Sentiment Analysis with Valency and Tendency Functions using BERT Sequence Model**

Along with existing devices (if needed). With this non-exclusive royalty-free right, Nusa Putra University has the right to store, transfer media/formats, process in the form of a database, maintain and publish my thesis as long as I keep my name as the author/creator and as the copyright owner.
This statement I made in truth.

Made in : Sukabumi
At the Date of : August 2024

That States
(Lutfil Khakim)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF LISTING

ABSTRACT

Sentiment Analysis with Valency and Tendency Functions using BERT Sequence Model

This study explores the potential of analyzing public opinion in Indonesia using advanced deep learning techniques to enhance sentiment analysis. Leveraging the BERT (Bidirectional Encoder Representations from Transformers) sequence model, specifically the BertForSequenceClassification model, we capture nuanced sentiment information through BERT's deep contextualized word representations. A dual-stage sentiment analysis framework is introduced, integrating the Tendency and Valency models to improve accuracy. The Tendency Model classifies texts into Low Tendency or High Tendency categories, while the Valency Model further refines sentiment analysis within the High Tendency data by evaluating sentiment intensity and distinguishing between positive and negative sentiments. This dual-stage approach significantly outperforms traditional single-stage methods, which achieved a lower accuracy of 35% due to their limited ability to capture nuanced sentiment variations. The dual-stage model demonstrates superior performance, achieving an accuracy of 82% and an F1 score of 78% on test data, indicating high precision in sentiment evaluation. The study highlights the effectiveness of combining deep learning techniques with a dual-stage framework to provide more accurate and contextually aware sentiment classification, advancing the analysis of public opinion with greater precision.

Keywords: Sentiment Analysis, Valency, Tendency, BERT, Deep Learning, Text Classification

CHAPTER I
INTRODUCTION

1.1 Background

In recent years, the use of social media has surged, embedding itself deeply within modern society. Internet users globally are no longer limited to using social media platforms merely to exchange personal information or communicate with friends, colleagues, and relatives. These platforms have evolved into spaces where users freely express their thoughts and opinions on a myriad of topics, ranging from products, people, events, trends, to social issues. This broad spectrum of topics reflects the diverse interests and concerns of users. Consequently, social media has become an essential tool for understanding public opinion.

Such expressions on social media provide invaluable insights for companies aiming to gauge customer sentiment towards their goods or services. The feedback collected from these platforms is instrumental in aiding business improvements. Companies analyze this feedback to refine their products and marketing strategies. This continuous loop of feedback and improvement helps businesses stay competitive. Understanding customer sentiment is thus a cornerstone of modern business strategy.

Governments, institutions, and public figures can leverage these insights to understand public perceptions and make informed decisions. By monitoring social media, they can gauge public response to policies and initiatives. This real-time feedback mechanism helps in shaping policies that are more aligned with public needs. Furthermore, it enhances transparency and trust between the public and governing bodies. Effective use of social media insights can thus lead to better governance.

From a consumer's perspective, shared opinions and sentiments about products on social media serve as crucial references in their decision-making processes before making purchases. Consumers trust reviews and feedback from other users more than traditional advertising. This exchange of information has transformed social media into a powerful tool for consumer awareness and education. It empowers consumers to make informed choices. Consequently, social media reviews have become a critical factor in consumer behavior.

The 2024 elections in Indonesia, particularly the presidential race, highlight the significant impact of social media on political discourse. Each debate, officially facilitated by the General Election Commission (KPU), ignites extensive discussions on social media. These discussions generate a wave of digital discourse, amplifying the reach of political messages. Statements made by presidential and vice-presidential candidates quickly become trending topics. This phenomenon underscores the power of social media in shaping electoral outcomes.

This dynamic interaction on social media underscores the pivotal role of these platforms in shaping public opinion. The polarizing effect it can have on society, especially during election periods, makes social media a crucial battleground in the political arena. Differences in opinions and the resulting societal polarization create a tense atmosphere. Social media becomes a reflection of the broader political climate. This emphasizes the importance of understanding and managing digital discourse.

The increasing influence of social networks in expressing opinions on current events and the rapid dissemination of online content has rendered online opinions a valuable asset for sentiment analysis. Online opinions, viewed as sentiments, reflect a person's views, opinions, or emotions. Consequently, the data generated on social media platforms has become a critical resource for sentiment analysis research. This data provides real-time insights into public sentiment. Researchers and analysts utilize this information to predict trends and behaviors.

Sentiment Analysis is a technique in natural language processing (NLP) utilized to identify, extract, or assess sentiments, opinions, or emotions contained within a text, such as product reviews, social media posts, or news articles (Jim et al., 2024; Nandwani & Verma, 2021). Visual Sentiment Analysis is a specific form of sentiment analysis that concentrates on visual data, such as images and videos, to recognize and evaluate the sentiments or emotions contained within that visual content (Bhoir & Jayamalini, 2021; H. Zhang et al., 2024).

Sentiment analysis is now widely recognized not only among researchers but also by businesses and governments (Birjali et al., 2021). Increased use of the internet has made the web a universal and most crucial source of information. Millions of individuals exchange their thoughts, ideas, expressions, emotions, and opinions on social media platforms such as Twitter, YouTube, Facebook, and others, adding a great deal of passion to human life nowadays (Srivastava & Kumar Soni, 2022). These opinions and sentiments are highly relevant to our daily lives, and as a result, there is a need to analyze this user-generated data to automatically monitor public opinion and assist in decision-making (Ramírez-Tinoco et al., 2018; Schuller et al., 2015). For example, Twitter posts have been used to predict election outcomes.

Sentiment analysis can be conducted at various levels, including document, sentence, and aspect levels. Document-level sentiment analysis determines the sentiment polarity of long texts such as news articles. Sentence-level sentiment analysis examines the sentiment of shorter texts, such as news headlines or social media comments on platforms like Twitter and Instagram. Aspect-level analysis focuses on specific components within a text. These various levels of sentiment analysis provide a comprehensive understanding of public sentiment across different contexts and text lengths.

1.2 Problem Statement

Public opinion in Indonesia holds significant potential as valuable information for sentiment analysis. The insights derived from analyzing public sentiments are crucial for evaluating the products and services of institutions or corporations. Public sentiment regarding any issue is typically categorized into positive, negative, or neutral responses. This classification helps in understanding the general mood of the population. It also assists in identifying areas needing attention or improvement.

Existing studies on sentiment analysis have aimed to detect subjective sentiment on specific topics, such as hotel reviews, mobile app reviews, and public opinion on current issues on platforms like Twitter (Aslan, 2023, 2023; Chaudhry et al., 2021; Mohbey et al., 2024). These studies predominantly use English-language data, reflecting a need for broader linguistic and cultural inclusivity in sentiment analysis research. Expanding research to include diverse languages will provide a more comprehensive understanding of global sentiments. This inclusivity is essential for accurate and relevant analysis. Broadening the scope of research can lead to more insightful and actionable findings.

The rapid growth of social media has led to an overwhelming amount of user-generated content, where individuals express their opinions on various topics. This proliferation of data presents a valuable opportunity for sentiment analysis to understand and predict public opinion. However, accurately capturing and interpreting these sentiments remains a significant challenge, particularly given the diverse linguistic and cultural contexts in which these opinions are expressed. Existing sentiment analysis research predominantly focuses on English-language data, resulting in a gap when it comes to understanding sentiments expressed in other languages, including Indonesian. This study seeks to address this gap by developing a robust sentiment analysis model that can effectively capture and interpret sentiments from Indonesian language.

To achieve this goal, the study leverages advanced deep learning techniques, specifically the BERT (Bidirectional Encoder Representations from Transformers) sequence model. BERT, built on the Transformer architecture introduced in the seminal paper "Attention Is All You Need," employs self-attention mechanisms to capture intricate relationships between words in a sentence (Vaswani et al., 2017). This bidirectional approach enables BERT to comprehend the context of each word based on its surrounding words, facilitating accurate sentiment analysis across diverse linguistic contexts. The tokenizer within BERT ensures comprehensive tokenization of text inputs, accommodating even out-of-vocabulary words.

In this study, we utilize the BertForSequenceClassification model, which is fine-tuned for sentiment analysis tasks. This model enhances the ability to capture nuanced sentiment information by leveraging BERT's deep contextualized word representations. Furthermore, the study enhances sentiment analysis by incorporating the concepts of valency and tendency into the BERT-based model. Valency represents the inherent positivity or negativity of individual words, while tendency captures the overall sentiment trajectory within a text. This dual approach aims to provide a comprehensive understanding of sentiments expressed in Indonesian social media discussions.

Additionally, contextual masking is introduced to define contexts such as individuals, political parties, or specific products, determining whether these contexts hold certain sentiments. This method, combined with the dual assessments of tendency and valency, aims to improve the accuracy of sentiment analysis by providing a more thorough evaluation. The outcomes of this research endeavor to offer actionable insights for businesses, governments, and institutions, empowering informed decision-making processes based on a deeper understanding of public sentiment dynamics. The research questions guiding this study are as follows:

a. How does the incorporation of contextual masking influence the accuracy of sentiment analysis in identifying sentiments related to specific entities such as individuals, political parties, and others?
b. What impact do the dual assessments of tendency and valency have on the overall accuracy and comprehensiveness of sentiment analysis in Indonesian language?
c. How does a dual-stage sentiment analysis model, utilizing both tendency and valency assessments, compare in performance to a single-stage model in accurately determining sentiment?

1.3 Research Objectives

a. To investigate the effect of contextual masking on the accuracy of sentiment analysis in identifying specific entities within Indonesian text data.
b. To evaluate the impact of dual assessments of tendency and valency on the accuracy and comprehensiveness of sentiment analysis.
c. To compare the performance of a dual-stage sentiment analysis model utilizing both tendency and valency assessments with a single-stage model in accurately determining sentiment.

1.4 Significance of Study

a. The integration of BERT (Bidirectional Encoder Representations from Transformers) for sequence classification provides robust contextual understanding, while the novel approach of dual-stage sentiment analysis offers comprehensive insights into public opinion. This combination can yield valuable information for businesses, policymakers, and researchers, facilitating more informed decision-making processes.

b. Most existing sentiment analysis research focuses on English-language data, leaving a gap in understanding sentiments expressed in other languages, including Indonesian. This study addresses this gap by developing and validating a model specifically tailored for Indonesian text, offering insights that can be applied across various sectors including business, politics, and public opinion research.

c. The outcomes of this research can be utilized in various practical applications, including market analysis, political sentiment tracking, and customer feedback evaluation. By accurately capturing public sentiment, organizations can tailor their strategies and responses to better align with public opinion and improve engagement.

d. The study contributes to the broader field of NLP by demonstrating the effectiveness of advanced deep learning techniques, such as BERT, in combination with innovative sentiment analysis frameworks. This can inspire further research and development in NLP, particularly in enhancing the capabilities of sentiment analysis models for diverse languages and contexts.

1.5 Limitation of Problems and Assumptions

a. One of the primary limitations of this study is the availability of labeled data. Sentiment analysis models, particularly those employing advanced techniques such as BERT and dual-stage sentiment analysis, rely heavily on large volumes of labeled data for training. The quality and quantity of labeled data directly impact the model's performance. In this study, the dataset might not encompass a sufficiently diverse range of contexts and sentiment expressions, leading to potential issues with overfitting and poor generalization. An insufficient amount of labeled data can hinder the model’s ability to accurately classify sentiments across various scenarios, affecting the overall reliability and robustness of the results.

b. While the dual-stage model aims to incorporate context into sentiment analysis, challenges remain in effectively understanding and integrating contextual nuances. Contextual understanding in sentiment analysis requires the model to grasp not only the sentiment conveyed but also the underlying context and subtleties of the text. Despite improvements, the model may still struggle with complex or ambiguous contexts, leading to potential inaccuracies in sentiment classification. The effectiveness of contextual analysis can vary based on the complexity of the language and the specific nuances of the data being analyzed.
c. The dual-stage model evaluates sentiment through valency and tendency components. However, accurately distinguishing between different sentiment intensities and contexts remains a challenge. The valency model, which assesses the positivity or negativity of sentiment, and the tendency model, which evaluates the sentiment's direction, may have limitations in capturing the full spectrum of sentiment expressions. In practice, these models might not always align perfectly with human judgment, particularly in complex or nuanced statements.
d. The study also addresses sentiment analysis in statements containing mixed sentiments, where both positive and negative sentiments are present. Accurately parsing and classifying mixed sentiments is inherently challenging and may lead to inconsistent results. The dual-stage model aims to improve handling of mixed sentiments, but there may still be limitations in effectively distinguishing and representing the various sentiment components within a single statement.
e. Another limitation is the generalizability of the model across different domains and contexts. The dataset used for training and evaluation may be specific to certain topics or styles of text, which can affect the model's performance when applied to other domains. The ability of the model to generalize and accurately classify sentiments in diverse contexts beyond the dataset is a crucial consideration for its practical applicability.
f. The study assumes that the dataset used for training and evaluation is representative of the broader range of sentiment expressions and contexts.

This assumption is critical for the validity of the sentiment analysis model's results. If the dataset does not adequately represent the diversity of sentiment expressions, the model's performance and generalizability could be adversely affected.

g. The study assumes the availability of adequate computational resources for training and evaluating the BERT sequence model and dual-stage sentiment analysis. The computational requirements for these models can be substantial, and the assumption of sufficient resources is necessary for conducting the analysis effectively.

CHAPTER V
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The research on "Sentiment as a Function of Valency and Tendency Using BERT Sequence Model" provides significant insights into enhancing sentiment analysis through advanced modeling techniques. This study leverages BERT (Bidirectional Encoder Representations from Transformers) for sequence classification, which employs Transformer-based architecture to process and understand text sequences. BERT's method of masking, where certain words in the text are hidden during training, enables the model to learn contextual relationships and nuances in language. This research distinguishes itself from traditional sentiment analysis methods by introducing a dual-stage approach that first evaluates sentiment in terms of tendency and valency, thus providing a more nuanced classification.

5.1.1 Model Performance

The dual-stage sentiment analysis system, integrating the Tendency and Valency models, demonstrates a comprehensive and effective approach to sentiment classification. The Tendency Model, evaluated on test data, achieved an accuracy of 0.81 and an F1 score of 0.78. This model adeptly classifies texts into Low Tendency or High Tendency categories, capturing the overall sentiment direction and providing a foundational understanding of sentiment orientation. By identifying whether a text exhibits a high or low tendency, the Tendency Model sets the stage for more nuanced sentiment analysis.

Following the initial classification, the Valency Model, which further refines sentiment analysis within the High Tendency data, achieved a higher accuracy of 0.83 and an F1 score of 0.82 on the test data. This model evaluates the sentiment intensity, distinguishing between positive and negative sentiments with high precision. The improved performance of the Valency Model highlights its effectiveness in assessing sentiment nuances within the context identified by the Tendency Model.

In contrast, traditional single-stage sentiment analysis methods, which do not incorporate separate assessments of tendency and valency, achieved a significantly lower accuracy of 0.35. This notable disparity underscores the limitations of single-stage approaches in capturing the nuanced and contextually variable nature of sentiments effectively.

The dual-stage system's overall performance on the test data, with an average accuracy of approximately 0.82 and a combined F1 score of 0.78, reflects its advanced capability in integrating tendency and valency for a more detailed and accurate sentiment analysis. While this accuracy indicates strong performance, it also suggests that the dual-stage system offers substantial improvements over traditional methods, demonstrating enhanced contextual awareness and precision in sentiment classification.

In summary, the integration of the Tendency and Valency Models represents a significant advancement in sentiment analysis, providing a robust and precise tool for capturing and differentiating nuanced sentiments effectively.

5.1.2 Strengths of the Dual-Stage Model

- Nuanced Sentiment Analysis: The dual-stage model's ability to evaluate both tendency and valency offers a more detailed understanding of sentiment, allowing for accurate differentiation between subtle sentiment variations. This capability is particularly useful in complex statements where context plays a crucial role.
- Contextual Awareness: The model's integration of contextual information enhances its ability to accurately reflect sentiment nuances. This is evident in scenarios where traditional context-free models may fail to capture the sentiment's true nature due to a lack of contextual understanding.

5.1.3 Limitations and Areas for Improvement:

- Dataset Limitations: Despite its strengths, the dual-stage model's accuracy is limited by the available dataset. The current dataset may not fully represent the range of possible contexts and sentiments, which affects the model's generalization and robustness. The accuracy achieved, while promising, highlights the need for a more extensive dataset to better capture diverse contexts and sentiment expressions.

- Handling Mixed Sentiments: The model struggles with accurately interpreting mixed sentiments within complex statements. This limitation underscores the need for further refinement in how the model processes and integrates multiple sentiment components.
- Need for Enhanced Labeling: The effectiveness of the model is also constrained by the quality and quantity of labeled data. More comprehensive labeling, particularly in varied contexts, will strengthen the model's ability to generalize and improve its predictive performance.

5.2 Recommendations

To enhance the effectiveness and applicability of sentiment analysis models, several key recommendations should be considered for future research and development.

Firstly, expanding and diversifying the dataset used for training and evaluation is crucial. A broader dataset that includes a variety of contexts, topics, and sentiment labels will enable the model to capture a wider range of sentiment expressions and nuances. This diversification will not only improve the model's ability to generalize but also enhance its robustness across different types of textual data. Incorporating more examples of mixed sentiments, complex sentence structures, and varying levels of emotional intensity will help the model to better understand and classify sentiment in diverse scenarios.

In addition to expanding the dataset, exploring alternative deep learning algorithms beyond BERT is recommended. While BERT has demonstrated significant capabilities in sequence classification, models such as GPT-4, RoBERTa, and XLNet offer different strengths and architectural innovations. For instance, GPT-4's capabilities in generating contextually rich text might provide complementary insights, while RoBERTa's enhancements over BERT could yield better performance in certain tasks. Experimenting with these models could reveal new ways to improve sentiment analysis, potentially leading to better performance in specific contexts or more nuanced sentiment classification.

Moreover, integrating advanced natural language processing (NLP) techniques can further refine sentiment analysis. Techniques such as syntactic parsing, which analyzes sentence structure, and semantic role labeling, which identifies roles of words in sentences, can add layers of contextual understanding. These methods can help the model to better interpret complex sentences and mixed sentiments, which are often challenging for traditional models to handle. For instance, parsing sentence structure can reveal underlying sentiments that are obscured by complex phrasing, while semantic role labeling can clarify the roles of different elements in a sentence, improving the accuracy of sentiment classification.

Additionally, continuous improvement of the dual-stage sentiment analysis system should be a priority. This includes refining the valency evaluation component to better distinguish between varying degrees of emotional intensity. Enhancing the model's sensitivity to different valency levels can lead to more precise sentiment analysis, especially in cases where the intensity of sentiment is crucial. Regular updates and training with new data will also help in adapting to evolving language use and sentiment expression patterns.

Finally, it is essential to ensure that evaluation and preprocessing methods are consistent throughout the model development lifecycle. Inconsistent methods between training and testing phases can lead to inaccurate results and hinder model performance. Establishing rigorous and uniform evaluation protocols will provide more reliable insights into the model's effectiveness and ensure that improvements are based on accurate assessments.

By addressing these recommendations, future research can advance sentiment analysis technologies, leading to more accurate, nuanced, and practical applications in real-world scenarios. These enhancements will contribute to the development of more robust sentiment analysis systems that are better equipped to handle the complexities of human language and emotion.

REFERENCE

A., V., & Sonawane, S. S. (2016). Sentiment Analysis of Twitter Data: A Survey of Techniques. *International Journal of Computer Applications*, *139*(11), 5–15. https://doi.org/10.5120/ijca2016908625

Ahmad, S. R., Bakar, A. A., & Yaakub, M. R. (2019). Ant colony optimization for text feature selection in sentiment analysis. *Intelligent Data Analysis*, *23*(1), 133–158. https://doi.org/10.3233/IDA-173740

Alqaryouti, O., Siyam, N., Abdel Monem, A., & Shaalan, K. (2024). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*, *20*(1/2), 142–161. https://doi.org/10.1016/j.aci.2019.11.003

Al-Smadi, M., Talafha, B., Al-Ayyoub, M., & Jararweh, Y. (2019). Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. *International Journal of Machine Learning and Cybernetics*, *10*(8), 2163–2175. https://doi.org/10.1007/s13042-018-0799-4

Ansari, G., Ahmad, T., & Doja, M. N. (2019). Hybrid Filter–Wrapper Feature Selection Method for Sentiment Classification. *Arabian Journal for Science and Engineering*, *44*(11), 9191–9208. https://doi.org/10.1007/s13369-019-04064-6

Armykav, R., Mantoro, T., Ayu, M. A., & Asian, J. (2023). Sentiment Analysis CNN Indonesia App Reviews on Play Store Using Naive Bayes Algorithm. *2023 International Conference on Technology, Engineering, and Computing Applications (ICTECA)*, 1–5. https://doi.org/10.1109/ICTECA60133.2023.10490854

Aslan, S. (2023). A deep learning-based sentiment analysis approach (MF-CNN-BILSTM) and topic modeling of tweets related to the Ukraine–Russia conflict. *Applied Soft Computing*, *143*, 110404. https://doi.org/10.1016/j.asoc.2023.110404

Berkovic, D., Ackerman, I. N., Briggs, A. M., & Ayton, D. (2020). Tweets by People With Arthritis During the COVID-19 Pandemic: Content and Sentiment Analysis. *Journal of Medical Internet Research*, *22*(12), e24550. https://doi.org/10.2196/24550

Bhoir, H., & Jayamalini, K. (2021). Visual Sentiment Analysis on Social Media Data. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 366–372. https://doi.org/10.32628/CSEIT2174101

Birjali, M., Beni-Hssane, A., & Erritali, M. (2017). Machine Learning and Semantic Sentiment Analysis based Algorithms for Suicide Sentiment Prediction in Social Networks. *Procedia Computer Science*, *113*, 65–72. https://doi.org/10.1016/j.procs.2017.08.290

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. https://doi.org/10.1016/j.knosys.2021.107134

Chaudhry, H. N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z. I., Shoaib, U., & Janjua, S. H. (2021). Sentiment Analysis of before and after Elections: Twitter Data of U.S. Election 2020. *Electronics*, *10*(17), 2082. https://doi.org/10.3390/electronics10172082

Chen, T., Samaranayake, P., Cen, X., Qi, M., & Lan, Y.-C. (2022). The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study. *Frontiers in Psychology*, *13*, 865702. https://doi.org/10.3389/fpsyg.2022.865702

Chen, T., Xu, R., He, Y., & Wang, X. (2017). Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications*, *72*, 221–230. https://doi.org/10.1016/j.eswa.2016.10.065

Chollet, F. (2018). *Deep learning with Python*. Manning Publications Co.

Claes, M., Farooq, U., Salman, I., Teern, A., Isomursu, M., & Halonen, R. (2024). Sentiment Analysis of Finnish Twitter Discussions on COVID-19 During the Pandemic. *SN Computer Science*, *5*(2), 266. https://doi.org/10.1007/s42979-023-02595-2

Conneau, A., Schwenk, H., Barrault, L., & Lecun, Y. (2016). *Very Deep Convolutional Networks for Text Classification* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1606.01781

Croucher, S. M., Nguyen, T., & Rahmani, D. (2020). Prejudice Toward Asian Americans in the Covid-19 Pandemic: The Effects of Social Media Use in the United States. *Frontiers in Communication*, *5*, 39. https://doi.org/10.3389/fcomm.2020.00039

Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. *Artificial Intelligence Review*, *56*(8), 8469–8510. https://doi.org/10.1007/s10462-022-10386-z

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1810.04805

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

D'Silva, J., & Sharma, U. (2022). Automatic text summarization of konkani texts using pre-trained word embeddings and deep learning. *International Journal of Electrical* *and Computer Engineering (IJECE)*, *12*(2), 1990. https://doi.org/10.11591/ijece.v12i2.pp1990-2000

Duric, A., & Song, F. (2012). Feature selection for sentiment analysis based on content and syntax models. *Decision Support Systems*, *53*(4), 704–711. https://doi.org/10.1016/j.dss.2012.05.023

Dzisevic, R., & Sesok, D. (2019). Text Classification using Different Feature Extraction Approaches. *2019 Open Conference of Electrical, Electronic and Information Sciences (eStream)*, 1–4. https://doi.org/10.1109/eStream.2019.8732167

Forman, G. (2003). *An Extensive Empirical Study of Feature Selection Metrics for Text Classification*. 1289–1305. https://doi.org/10.5555/944919.944974

Garg, S. B., & Subrahmanyam, V. V. (2022). Sentiment Analysis: Choosing the Right Word Embedding for Deep Learning Model. In M. Bianchini, V. Piuri, S. Das, & R. N. Shaw (Eds.), *Advanced Computing and Intelligent Technologies* (Vol. 218, pp. 417–428). Springer Singapore. https://doi.org/10.1007/978-981-16-2164-2_33

Gunawan, O. A. C., Ginting, D. A., Handoyo, R. A., Willy, A., Chandra, S., Purnomo, F., & Yulistiani, R. (2022). Sentiment Analysis towards Face-to-Face School Activities During the COVID-19 pandemic in Indonesia. *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*, 1–7. https://doi.org/10.1109/ICORIS56080.2022.10031482

Hu, Y., Ding, J., Dou, Z., & Chang, H. (2022). Short-Text Classification Detector: A Bert-Based Mental Approach. *Computational Intelligence and Neuroscience*, *2022*, 1–11. https://doi.org/10.1155/2022/8660828

Hur, M., Kang, P., & Cho, S. (2016). Box-office forecasting based on sentiments of movie reviews and Independent subspace method. *Information Sciences*, *372*, 608–624. https://doi.org/10.1016/j.ins.2016.08.027

Iscus, F., & Girsang, A. S. (2022). Sentiment Analysis of COVID-19 Public Activity Restriction (PPKM) Impact using BERT Method. *International Journal of Engineering Trends and Technology*, *70*(12), 281–288. https://doi.org/10.14445/22315381/IJETT-V70I12P226

Jayadianti, H., Kaswidjanti, W., Utomo, A. T., Saifullah, S., Dwiyanto, F. A., & Drezewski, R. (2022). Sentiment analysis of Indonesian reviews using fine-tuning IndoBERT and R-CNN. *ILKOM Jurnal Ilmiah*, *14*(3), 348–354. https://doi.org/10.33096/ilkom.v14i3.1505.348-354

Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. F. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, *6*, 100059. https://doi.org/10.1016/j.nlp.2024.100059

Kasri, M., Birjali, M., & Beni-Hssane, A. (2019). A comparison of features extraction methods for Arabic sentiment analysis. *Proceedings of the 4th International Conference on Big Data and Internet of Things*, 1–6. https://doi.org/10.1145/3372938.3372998

Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746–1751. https://doi.org/10.3115/v1/D14-1181

Kou, G., Yang, P., Peng, Y., Xiao, F., Chen, Y., & Alsaadi, F. E. (2020). Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. *Applied Soft Computing*, *86*, 105836. https://doi.org/10.1016/j.asoc.2019.105836

Kumar, R., & Kaur, J. (2020). Random Forest-Based Sarcastic Tweet Classification Using Multiple Feature Collection. In S. Tanwar, S. Tyagi, & N. Kumar (Eds.), *Multimedia*

Big Data Computing for IoT Applications (Vol. 163, pp. 131–160). Springer Singapore. https://doi.org/10.1007/978-981-13-8759-3_5

Liang, H., Sun, X., Sun, Y., & Gao, Y. (2017). Text feature extraction based on deep learning: A review. *EURASIP Journal on Wireless Communications and Networking*, *2017*(1), 211. https://doi.org/10.1186/s13638-017-0993-1

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Springer International Publishing. https://doi.org/10.1007/978-3-031-02145-9

Liu, B., & Zhang, L. (2012). A Survey of Opinion Mining and Sentiment Analysis. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 415–463). Springer US. https://doi.org/10.1007/978-1-4614-3223-4_13

Liu, H., Zhou, M., & Liu, Q. (2019). An embedded feature selection method for imbalanced data classification. *IEEE/CAA Journal of Automatica Sinica*, *6*(3), 703–715. https://doi.org/10.1109/JAS.2019.1911447

Liu, Y., Liu, B., Shan, L., & Wang, X. (2018). Modelling context with neural networks for recommending idioms in essay writing. *Neurocomputing*, *275*, 2287–2293. https://doi.org/10.1016/j.neucom.2017.11.005

Lohith, C., Chandramouli, H., Balasingam, U., & Arun Kumar, S. (2023). Aspect Oriented Sentiment Analysis on Customer Reviews on Restaurant Using the LDA and BERT Method. *SN Computer Science*, *4*(4), 399. https://doi.org/10.1007/s42979-022-01634-8

Loia, V., & Senatore, S. (2014). A fuzzy-oriented sentic analysis to capture the human emotion in Web-based content. *Knowledge-Based Systems*, *58*, 75–85. https://doi.org/10.1016/j.knosys.2013.09.024

Ma, S., Yang, P., Liu, T., Li, P., Zhou, J., & Sun, X. (2019). Key Fact as Pivot: A Two-Stage Model for Low Resource Table-to-Text Generation. *Proceedings of the 57th*

Annual Meeting of the Association for Computational Linguistics, 2047–2057. https://doi.org/10.18653/v1/P19-1197

Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., & Potts, C. (2011). *Learning Word Vectors for Sentiment Analysis*. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, 142–150.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1310.4546

Mohbey, K. K., Meena, G., Kumar, S., & Lokesh, K. (2024). A CNN-LSTM-Based Hybrid Deep Learning Approach for Sentiment Analysis on Monkeypox Tweets. *New Generation Computing*, *42*(1), 89–107. https://doi.org/10.1007/s00354-023-00227-0

Mowlaei, M. E., Saniee Abadeh, M., & Keshavarz, H. (2020). Aspect-based sentiment analysis using adaptive aspect-based lexicons. *Expert Systems with Applications*, *148*, 113234. https://doi.org/10.1016/j.eswa.2020.113234

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Mutinda, J., Mwangi, W., & Okeyo, G. (2021). Lexicon-pointed hybrid N-gram Features Extraction Model (LeNFEM) for sentence level sentiment analysis. *Engineering Reports*, *3*(8), e12374. https://doi.org/10.1002/eng2.12374

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 81. https://doi.org/10.1007/s13278-021-00776-6

Nissa, N. K., & Yulianti, E. (2023). Multi-label text classification of Indonesian customer reviews using bidirectional encoder representations from transformers language model. *International Journal of Electrical and Computer Engineering (IJECE)*, *13*(5), 5641. https://doi.org/10.11591/ijece.v13i5.pp5641-5652

Nurdeni, D. A., Budi, I., & Santoso, A. B. (2021). Sentiment Analysis on Covid19 Vaccines in Indonesia: From The Perspective of Sinovac and Pfizer. *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, 122–127. https://doi.org/10.1109/EIConCIT50028.2021.9431852

Park, B., Jang, I. S., & Kwak, D. (2024). Sentiment analysis of the COVID-19 vaccine perception. *Health Informatics Journal*, *30*(1), 14604582241236131. https://doi.org/10.1177/14604582241236131

Patel, A., Oza, P., & Agrawal, S. (2023). Sentiment Analysis of Customer Feedback and Reviews for Airline Services using Language Representation Model. *Procedia Computer Science*, *218*, 2459–2467. https://doi.org/10.1016/j.procs.2023.01.221

Polanyi, L., & Zaenen, A. (2006). Contextual Valence Shifters. In J. G. Shanahan, Y. Qu, & J. Wiebe (Eds.), *Computing Attitude and Affect in Text: Theory and Applications* (Vol. 20, pp. 1–10). Springer-Verlag. https://doi.org/10.1007/1-4020-4102-0_1

Pristiyono, Ritonga, M., Ihsan, M. A. A., Anjar, A., & Rambe, F. H. (2021). Sentiment analysis of COVID-19 vaccine in Indonesia using Naïve Bayes Algorithm. *IOP Conference Series: Materials Science and Engineering*, *1088*(1), 012045. https://doi.org/10.1088/1757-899X/1088/1/012045

Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning. *Sensors*, *22*(11), 4157. https://doi.org/10.3390/s22114157

Ramírez-Tinoco, F. J., Alor-Hernández, G., Sánchez-Cervantes, J. L., Olivares-Zepahua, B. A., & Rodríguez-Mazahua, L. (2018). A Brief Review on the Use of Sentiment Analysis Approaches in Social Networks. In J. Mejia, M. Muñoz, Á. Rocha, Y. Quiñonez, & J. Calvo-Manzano (Eds.), *Trends and Applications in Software Engineering* (Vol. 688, pp. 263–273). Springer International Publishing. https://doi.org/10.1007/978-3-319-69341-5_24

Romadhony, A., Al Faraby, S., Rismala, R., Wisesty, U. N., & Arifianto, A. (2024). Sentiment Analysis on a Large Indonesian Product Review Dataset. *Journal of Information Systems Engineering and Business Intelligence*, *10*(1), 167–178. https://doi.org/10.20473/jisebi.10.1.167-178

Sánchez-Maroño, N., Alonso-Betanzos, A., & Calvo-Estévez, R. M. (2009). A Wrapper Method for Feature Selection in Multiple Classes Datasets. In J. Cabestany, F. Sandoval, A. Prieto, & J. M. Corchado (Eds.), *Bio-Inspired Systems: Computational and Ambient Intelligence* (Vol. 5517, pp. 456–463). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02478-8_57

Schuller, B., Mousa, A. E., & Vryniotis, V. (2015). Sentiment analysis and opinion mining: On optimal parameters and performances. *WIREs Data Mining and Knowledge Discovery*, *5*(5), 255–263. https://doi.org/10.1002/widm.1159

Simanjuntak, L. F., Mahendra, R., & Yulianti, E. (2022). We Know You Are Living in Bali: Location Prediction of Twitter Users Using BERT Language Model. *Big Data and Cognitive Computing*, *6*(3), 77. https://doi.org/10.3390/bdcc6030077

Singh, T., & Kumari, M. (2016). Role of Text Pre-processing in Twitter Sentiment Analysis. *Procedia Computer Science*, *89*, 549–554. https://doi.org/10.1016/j.procs.2016.06.095

Srivastava, D., & Kumar Soni, V. (2022). A Systematic Review On Sentiment Analysis Approaches. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 01–06. https://doi.org/10.1109/ICACITE53722.2022.9823769

Tetko, I. V., Karpov, P., Van Deursen, R., & Godin, G. (2020). State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis. *Nature Communications*, *11*(1), 5575. https://doi.org/10.1038/s41467-020-19266-y

Tubishat, M., Idris, N., & Abushariah, M. A. M. (2018). Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges. *Information Processing & Management*, *54*(4), 545–563. https://doi.org/10.1016/j.ipm.2018.03.008

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (Version 7). arXiv. https://doi.org/10.48550/ARXIV.1706.03762

Venugopalan, M., & Gupta, D. (2015). Exploring sentiment analysis on twitter data. *2015 Eighth International Conference on Contemporary Computing (IC3)*, 241–247. https://doi.org/10.1109/IC3.2015.7346686

Wan, Y., & Gao, Q. (2015). An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis. *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, 1318–1325. https://doi.org/10.1109/ICDMW.2015.7

Wang, S., & Manning, C. (2012). *Baselines and Bigrams: Simple, Good Sentiment and Topic Classification*. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 90–94.

Wang, Y., Huang, M., Zhu, X., & Zhao, L. (2016). Attention-based LSTM for Aspect-level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 606–615. https://doi.org/10.18653/v1/D16-1058

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Wirawan, C. (2021). *Indonesian BERT base model (uncased)* [Dataset].

Yang, P., & Chen, Y. (2017). A survey on sentiment analysis by using machine learning methods. *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 117–121. https://doi.org/10.1109/ITNEC.2017.8284920

Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016). Hierarchical Attention Networks for Document Classification. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1480–1489. https://doi.org/10.18653/v1/N16-1174

Yu, J., Jiang, J., & Xia, R. (2019). Global Inference for Aspect and Opinion Terms Co-Extraction Based on Multi-Task Neural Networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *27*(1), 168–177. https://doi.org/10.1109/TASLP.2018.2875170

Zhang, D., & Lee, W. S. (2003). Question classification using support vector machines. *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval*, 26–32. https://doi.org/10.1145/860435.860443

Zhang, H., Liu, Y., Xiong, Z., Wu, Z., & Xu, D. (2024). Visual sentiment analysis with semantic correlation enhancement. *Complex & Intelligent Systems*, *10*(2), 2869–2881. https://doi.org/10.1007/s40747-023-01296-w

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *WIREs Data Mining and Knowledge Discovery*, *8*(4), e1253. https://doi.org/10.1002/widm.1253

Zhang, X., Zhao, J., & LeCun, Y. (2015). *Character-level Convolutional Networks for Text Classification* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1509.01626

Zhao, Z., Rao, G., & Feng, Z. (2017a). DFDS: A Domain-Independent Framework for Document-Level Sentiment Analysis Based on RST. In L. Chen, C. S. Jensen, C. Shahabi, X. Yang, & X. Lian (Eds.), *Web and Big Data* (Vol. 10366, pp. 297–310). Springer International Publishing. https://doi.org/10.1007/978-3-319-63579-8_23

Zhao, Z., Rao, G., & Feng, Z. (2017b). DFDS: A Domain-Independent Framework for Document-Level Sentiment Analysis Based on RST. In L. Chen, C. S. Jensen, C. Shahabi, X. Yang, & X. Lian (Eds.), *Web and Big Data* (Vol. 10366, pp. 297–310). Springer International Publishing. https://doi.org/10.1007/978-3-319-63579-8_23