TEXT EMOTION ANALYSIS USING DEEP LEARNING ON NATIONAL ELECTION

A THESIS

A Partial Requirement to Fulfill for Master Degree in Computer Science

ALFIN ANDROMEDA

20211030032

Supervisors:

Prof. Ir.Teddy Mantoro, MSc., PhD., SMIEEE

Rahmadya Trias Handayanto, S.T., M.Kom., PhD

SCHOOL OF COMPUTER SCIENCE

NUSA PUTRA UNIVERSITY

STATEMENT OF AUTENTICITY

The undersigned below:

Name	:	Alfin Andromeda
ID of student	:	20210130032
Faculty	:	Computer Science
The Tittle of Thesis	:	**TEXT EMOTION ANALYSIS USING DEEP LEARNING ON NATIONAL ELECTION**

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, July 2024

Writer,

Materai 10000

Alfin Andromeda

20210130032

APPROVAL OF THESIS

The Tittle of Thesis	:	TEXT EMOTION ANALYSIS USING DEEP LEARNING ON NATIONAL ELECTION
Name	:	Alfin Andromeda
ID of student	:	20210130032

This thesis has been reviewed and approved

Sukabumi, July 2024

Head of Study Program,	Supervisor,
Prof. Ir.Teddy Mantoro, MSc., PhD., SMIEEE	Prof. Ir.Teddy Mantoro, MSc., PhD., SMIEEE
NIDN.	NIDN.

Nusa Putra UNIVERSITY 2002 AMOR DEUS, PARENTIUM, CONCERVIS

THESIS APPROVAL

The Tittle of Thesis	:	TEXT EMOTION ANALYSIS USING DEEP LEARNING ON NATIONAL ELECTION
Name	:	Alfin Andromeda
ID of student	:	20210130032

This Thesis has been tested and defended in front of the Board of Examiners in Thesis session on January 14, 2024. In our review, this Thesis adequate in terms of quality for the purpose of awarding the Master of Computer Degree.

Sukabumi, January 2024

Supervisor 1,	Examiner 1,
Prof. Ir.Teddy Mantoro, MSc., PhD., SMIEEE	Prof. Ir. Media Anugera Ayu, M.Sc., PhD
Supervisor 2,	Examiner 2,
Rahmadya Trias Handayanto, S.T., M.Kom., PhD.	Umar Aditiawarman, S.T., M.Sc., PhD

Nusa Putra UNIVERSITY 2002 AMOR DEUS, PARENTIUM, CONCERVIS

PUBLICATION APPROVAL

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

Name	:	Alfin Andromeda
ID of student	:	20210130032
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: **Text Emotion Analysis Using Deep Learning On National Election**. 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.

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS

Made in : Sukabumi

At the Date of : July 2024

That States

(Alfin Andromeda)

TABLE OF CONTENTS

STATEMENT OF AUTENTICITY ..... *ii*

APPROVAL OF THESIS ..... *iii*

THESIS APPROVAL ..... *iv*

PUBLICATION APPROVAL ..... *v*

TABLE OF CONTENTS ..... *vi*

LIST OF TABLES ..... *ix*

LIST OF FIGURES ..... *x*

FOREWORD ..... *xi*

ABSTRACT ..... *xii*

CHAPTER I INTRODUCTION *1*

1.1 Research Background **1**

1.2 Problem Statement **3**

1.3 Research Objectives **3**

1.4 Significance of Research **4**

1.5 Limitation of Problems and Assumptions **4**

1.6 Thesis Structure **4**

CHAPTER II LITERATURE REVIEW *5*

2.1 Literature Review **5**

2.1.1 Related Work 5

2.1.2 Social Media Impact on Election 9

2.1.3 Basic Emotion Theory 11

2.1.4 BERT Method 12

2.2 Research Flow **15**

Nusa Putra UNIVERSITY 2002 AMOR DEUS, PARENTIUM, CONCERVIS

CHAPTER III RESEARCH METHODOLOGY *17*

3.1 Study Approach Method **17**

3.2 Research Process **17**

3.3 Data Collection **18**

3.4 Data Pre-Processing **19**

3.5 Model Training **20**

3.5.1 Data Preparation 20

3.5.2 Model and Data Loader Preparation 21

3.5.3 Training the Model 21

3.5.4 Model Testing Technique 22

CHAPTER IV RESEARCH RESULT AND DISCUSSION *25*

4.1 Research Result **25**

4.1.1 Dataset Training Acquisition 25

4.1.2 Dataset Twitter Acquisition 26

4.1.3 Preprocessing data 29

4.1.4 BERT Model Training 35

4.1.5 Model Evaluation 41

4.1.6 Emotion Recognition on Election Data 45

4.2 Discussion **48**

4.2.1 The Using BERT for Emotion Recognition Model 48

4.2.2 Emotion Analysis on Indonesia Election 2024 49

4.2.3 Advanced Exploration on Emotion Analysis ..... **Error! Bookmark not defined.**

CHAPTER V CONCLUSIONS AND RECOMMENDATIONS *58*

5.1 Conclusions **58**

5.2 Recommendations **58**

REFERENCES *60*

Nusa Putra UNIVERSITY
AMOR DEUS, PARENTIUM, CONCERVIS

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS

LIST OF TABLES

Table 2.1.2 Related Works 9
Table 4.1.1 Tabel of Dataset for Training the Model 25
Table 4.1.2 Illustration of Lowercasing 31
Table 4.1.3 Illustration of URL Cleansing 32
Table 4.1.4 Illustration of URL Cleansing 32
Table 4.1.5 Illustration of Erasing Stopwords and Stemming 33
Table 4.1.6 Illustration of Slang Handling 35

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS

LIST OF FIGURES

Figure 2.1.1 Overall pre-training and fine-tuning procedures for BERT 13
Figure 2.1.2 Illustrations of Fine-tuning BERT Sentence Pair.....14
Figure 2.2.1 Experiment Process.....15
Figure 3.2.1 Research Process.....17
Figure 3.5.1 Example of Confusion Matrix.....23
Figure 4.1.1 Code for defining credential API and access Token for Twitter 26
Figure 4.1.2 Code for acquistion Public Tweet.....26
Figure 4.1.3 Twitter Dataset on Election Topic.....27
Figure 4.1.4 Pseudocode for Preprocessing Training Datasets 29
Figure 4.1.5 Pseudocode for Preprocessing Twitter/X Data.....30
Figure 4.1.6 Pseudocode for Training Model Preparation.....35
Figure 4.1.7 Pseudocode for Implementing BERT on Pytorch 37
Figure 4.1.8 Pseudocode for training the model.....39
Figure 4.1.9 Result 70% Data Training : 30% Data Testing 42
Figure 4.1.10 Result 75% Data Training : 25% Data Testing 42
Figure 4.1.11 Result 80% Data Training : 20% Data Testing 42
Figure 4.1.12 Graph Training Loss per Epoch and Accuracy per Epoch 44
Figure 4.1.13 Code for Processing Election Data 46
Figure 4.1.14 Frequency of Each Emotion on Election Dataset 47
Figure 4.2.1 Radar Diagram of Each Emotion 51
Figure 4.2.2 Scatter Plot of Frequency by Emotion Tendency.....52

FOREWORD

I am immensely grateful to present my thesis titled "Text Emotion Analysis using Deep Learning on National Elections." This work represents not just an academic pursuit but a journey into the complex interplay of technology, emotions, and political discourse.

Firstly, I extend my deepest gratitude to my supervisors, Prof. Ir. Teddy Mantoro, MSc., PhD., SMIEEE, and Rahmadya Trias Handayanto, S.T., M.Kom., PhD. Their expertise, guidance, and unwavering support have been invaluable to this research. I also express my sincere thanks to the faculty members and staff at Nusa Putra University. Their encouragement and academic rigor have provided a fertile ground for learning and exploration. I am also grateful to my peers and fellow researchers at the university, whose intellectual companionship and collaborative spirit have been both inspiring and enriching.

I owe a debt of gratitude to my family and friends for their endless support and patience. Their unwavering belief in my capabilities and constant encouragement have been a source of strength and motivation throughout this journey.

This thesis aims to delve into the emotional narratives present in Twitter discussions about Indonesia's national elections, employing deep learning techniques. The aim is to not only harness computational power but also to gain insights into the nation's emotional response during this significant democratic process. Lastly, I am grateful for the open-source resources and data that have been integral to this research. The open-source community's contributions have been fundamental in making this study possible.

Presenting this thesis is a milestone in a journey of learning and dedication. I hope that this work contributes meaningfully to the fields of emotion analysis and deep learning and sheds light on how technology can intersect with political and social realms.

Alfin Andromeda

Master's Student, Nusa Putra University

ABSTRACT

Indonesia's digital landscape, notably social media, is increasingly influencing politics, necessitating sophisticated analysis methods to comprehend public sentiment. However, traditional sentiment categorization is often biased, underscoring the need for nuanced emotion analysis. This study addresses this by using deep learning algorithms for emotion analysis on national elections data, aiming to accurately understand public emotions. This study began with a complex data collection phase, using manual downloading and data crawling techniques to gather tweets from Twitter. This was followed by extensive preprocessing, which involved case folding, cleansing, tokenizing, handling slang, and stemming to prepare the data for further stages. The refined data was split for modeling and analysis, with the BERT algorithm applied for training and subsequent evaluation. Ultimately, the most accurate model was used for emotion recognition on the election dataset, facilitating a comprehensive data exploration and analysis. The study resulted in a proficient use of the BERT model for emotion classification, revealing its capability for comprehensive emotion detection in varied text data. Additionally, the emotion analysis executed on an election-related Twitter dataset exhibited a polarity in the emotional landscape, with strong emotions such as "Anger" and "Love" dominating, thus mirroring diverse public opinions towards election events.

Keyword: BERT, Emotion Analysis, National Election, Social Media

CHAPTER I
INTRODUCTION

1.1 Research Background

The advancement of the digital era and technological progress enabled Indonesians to easily obtain information. According to a report by Statista, there was an increase in the number of internet users in Indonesia. By 2022, the number of internet users in Indonesia had reached around 224 million, with a projected increase to approximately 270 million by 2028, according to this report (*Indonesia: Number of Internet Users 2028 | Statista*, n.d.).

The rise in internet users and the growing affordability of smartphones effectively transformed the social media landscape in Indonesia. A significant portion of the population owned smartphones and utilized them extensively to simplify various aspects of their lives. For many Indonesians, social media served as a convenient means of connecting with family members residing in remote areas across the archipelago, facilitated constant interaction with friends, and kept them informed about daily news updates (*Indonesia: Social Network Penetration 2022 | Statista*, n.d.).

In recent times, social media has emerged as a valuable advertising and marketing tool for Indonesian brands. Marketing strategies on social media, such as influencer marketing and social commerce, are deemed crucial in a country where the average daily time spent on social media reaches three hours and twenty minutes. Brands aim to expand their customer base and enhance brand awareness by actively participating on social media platforms or collaborating with influencers. According to a recent study, 68 percent of Indonesians follow one to four influencers on social media. Moreover, approximately 80 percent of social media users in the country expressed their inclination to seek further information about a brand or product endorsed by influencers when exposed to influencer marketing content (*Social Media in Indonesia - Statistics & Facts | Statista*, n.d.).

Social media has gained immense popularity as a platform for receiving and engaging with political information. This has led to increased attention

from political marketers, responsible for political campaigning, due to the uptake of political discourse on social media (Weismueller et al., 2022). Southeast Asia is at the forefront of global discussions regarding the influence of social media campaigns on society and politics (Tapsell, 2021). In terms of campaign tactics, social media campaigners in the region have pioneered various innovative approaches, such as employing paid workers to promote certain narratives and attack critics on Twitter, utilizing Facebook for advertising and candidate promotion, and coordinating the dissemination of material through closed messenger groups like WhatsApp. While social media still serves as a platform for direct communication between politicians and citizens through their personal Facebook, Twitter, and Instagram accounts, it is just a part of a larger network in which these platforms shape political narratives online. The increasingly clandestine nature of social media campaigning has led to the creation of content that aims to shift election discourse towards identity politics, provoke outrage, and generate fictional and conspiratorial narratives.

One of the challenges posed by social media is the dissemination of fake news or the conduct of black campaigns through its platforms. According to CNN Indonesia's report, in 2018, there were instances on Twitter where false information circulated. For example, a Twitter user made an unfounded claim that Khofifah Indar Parwansa, a candidate for East Java governor, intended to legalize prostitution if elected. Another Twitter user similarly alleged that Uu Ruzhanul Ulum, a candidate for vice governor of West Java, had the potential to engage in corruption due to having four wives (Pratama & Rakhmawati, 2019).

Emotions play a fundamental role in the decision-making processes of individuals on a daily basis (Kratzwald et al., 2018). Moreover, emotional states have a subtle impact on human communication, attention, and the ability to remember information. While humans possess a natural ability to recognize and interpret emotions, computational algorithms face significant challenges in performing these tasks. This has led to the development of affective computing, which encompasses techniques aimed at detecting, recognizing, and predicting

human emotions such as joy, anger, sadness, trust, surprise, and anticipation (Balakrishnan & Kaur, 2019). The ultimate objective is to adapt computational systems to these emotional states. As a result, these computer systems not only demonstrate empathy but also offer tailored decision support based on individuals' emotional states.

Emotion analysis plays a crucial role in political issues (Huddy, 2015). It helps in understanding the sentiments and reactions of the public towards political events, policies, and leaders. By analyzing emotions expressed on social media, surveys, and public forums, policymakers can gain valuable insights into the concerns, frustrations, and expectations of the people. This information enables them to make informed decisions, tailor their communication strategies, and address the emotional needs of the public effectively. Emotion analysis also aids in identifying potential shifts in public sentiment, allowing politicians to adapt their approach and engage with the electorate in a more meaningful way. Ultimately, understanding and incorporating emotion analysis in political discourse can lead to better representation and governance

1.2 Problem Statement

The current trend in text analysis focusing on categorizing sentiments as positive or negative for understanding public opinion on elections tends to be biased due to the inherent subjectivity involved.

Incorporating emotion analysis can provide a more nuanced and accurate understanding of prevailing opinions and emotions within society regarding the upcoming 2024 elections.

This study aims to apply deep learning algorithms to develop emotion analysis, focusing on the topic of the 2024 elections.

1.3 Research Objectives

1. To explore the application of deep learning techniques in detecting and analyzing the emotional content of textual data.
2. To delve into the dynamics of public emotions to illustrate the societal situation leading up to the 2024 elections.

1.4 Significance of Research

1. This study is significant as it leverages deep learning algorithms to enhance the detection and analysis of emotions in Indonesian texts, thereby advancing sentiment analysis capabilities for the Indonesian language.
2. By revealing the emotional dynamics from Twitter users' tweets, this research provides valuable insights into the public sentiment leading up to the 2024 elections. These insights can help policymakers and political campaigners tailor their strategies, ensuring a more informed, responsive, and inclusive electoral process.

1.5 Limitation of Problems and Assumptions

The scope of this research focuses on implementing the deep learning algorithm BERT (Bidirectional Encoder Representations from Transformers) to identify five basic emotions (fear, joy, sadness, trust, anger) in the Indonesian language. The model will be evaluated for its accuracy and applied to analyze two datasets of tweets from the social media platform Twitter. The first dataset was collected during the period of March 1, 2023, to March 30, 2023, and the second dataset was collected during the period of January 1, 2024, to January 31, 2024, both containing the keyword "Pemilu" (election in Bahasa Indonesia).

1.6 Thesis Structure

The rest of thesis is organized as follows:

1. Chapter I describes the background of problem that will be discussed in the thesis.
2. Chapter II describes the literature review of thesis.
3. Chapter III describes the methodology of thesis.
4. Chapter IV presents the expereiment result and discussion.
5. Chapter V presents the conclusion the thesis and future work.

Nusa Putra
AMOR DEUS, PARENTIUM, CONCERVIS

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS

CHAPTER V
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

Based on the results and findings from this study, the following conclusions are:

1. This research indicates that the application of BERT models for emotion classification tasks can achieve excellent results, particularly in the realms of accuracy, precision, recall, and F1-score. The model exhibited substantial learning capabilities over the course of the training epochs, as evident by the reduction in training loss and increase in performance metrics. The effective differentiation of complex emotional states such as anger, fear, sad, joy, and love is testament to the intricate learning potential of the BERT model, opening up possibilities for nuanced emotion detection in diverse text data.
2. This study successfully carried out an emotion analysis on a Twitter dataset related to the election. The use of BERT, a sophisticated deep learning model, has shown to be quite effective in understanding and classifying emotions from textual data. The results reveal a considerable dominance of strong emotions such as "Anger" and "Love", indicating a highly polarized emotional landscape in the discourse surrounding the election. These emotions, reflected in tweets, represent diverse opinions and attitudes towards various election-related events and candidates. The prevalence of "Anger" might be an indication of widespread discontent or frustration, whereas the substantial representation of "Love" could suggest strong support or endorsement for certain candidates or causes.

5.2 Recommendations

Based on the results and findings from this study, the following recommendation for further research are proposed:

1. The current model relies solely on the textual content of the tweets for emotion classification. However, in social media discourse, the context often plays a crucial role in determining the sentiment or emotion of a statement. This could include information like the profile of the tweeter (e.g., political affiliations, geographical location), the wider conversational

thread in which the tweet is posted, or even temporal context (e.g., in relation to specific events or times). Incorporating such contextual information could make the model more sensitive to the intricacies of social media discourse and potentially enhance its emotion classification performance. To accomplish this, advanced NLP techniques like attention mechanisms or transformer models could be utilized to effectively capture and incorporate this contextual information into the model.

2. While this study focused on five primary emotions, it is recommended that future research should consider extending the spectrum of emotions. Human emotional responses can be incredibly nuanced and diverse, and this study's categories might not capture the full extent of these emotions. A wider range of emotion categories, potentially including secondary emotions or more complex emotional states, can provide more granular insights into public sentiment. This would require the development of more complex and nuanced training datasets, and could potentially involve improvements in the BERT model to more effectively classify these more diverse emotions.

Nusa Putra
2002
AMOR DEUS, PARENTIUM, CONCERVIS

REFERENCES

Anggraheni, P., Setyowati, N. T., & Harry, H. (2021). Social Media And Political Participation In Indonesia: Restrictions Access At Announcement Results Of 2019 Presidential Election. *Aspiration Journal*, *2*(1), 85–123. https://doi.org/10.56353/aspiration.v2i1.23

Antypas, D., Preece, A., & Camacho-Collados, J. (2023). Negativity spreads faster: A large-scale multilingual twitter analysis on the role of sentiment in political communication. *Online Social Networks and Media*, *33*. https://doi.org/10.1016/j.osnem.2023.100242

Balakrishnan, V., & Kaur, W. (2019). String-based multinomial naïve bayes for emotion detection among facebook diabetes community. *Procedia Computer Science*, *159*, 30–37. https://doi.org/10.1016/j.procs.2019.09.157

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. 4171–4186. http://arxiv.org/abs/1810.04805

Huddy, L. (2015). Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity. *American Political Science Review*, *109*(1). https://doi.org/10.1017/S0003055414000604

Indonesia: number of internet users 2028 | Statista. (n.d.). Retrieved July 9, 2023, from https://www.statista.com/statistics/254456/number-of-internet-users-in-indonesia/

Indonesia: social network penetration 2022 | Statista. (n.d.). Retrieved July 9, 2023, from https://www.statista.com/statistics/284437/indonesia-social-network-penetration/

Kratzwald, B., Ilić, S., Kraus, M., Feuerriegel, S., & Prendinger, H. (2018). *Deep learning for affective computing: Text-based emotion recognition in decision support*. https://doi.org/10.1016/j.dss.2018.09.002

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lin, S. Y., Kung, Y. C., & Leu, F. Y. (2022). Predictive intelligence in harmful news identification by BERT-based ensemble learning model with text sentiment

Nusa Putra UNIVERSITY

analysis. *Information Processing and Management*, *59*(2). https://doi.org/10.1016/j.ipm.2022.102872

Nabiilah, G. Z., Prasetyo, S. Y., Izdihar, Z. N., & Girsang, A. S. (2023). BERT base model for toxic comment analysis on Indonesian social media. *Procedia Computer Science*, *216*, 714–721. https://doi.org/10.1016/j.procs.2022.12.188

Parrott, W. G. (2001). *Emotions in social psychology: Essential readings.* psychology press.

Pratama, P. G., & Rakhmawati, N. A. (2019). Social bot detection on 2019 Indonesia president candidate's supporter's tweets. *Procedia Computer Science*, *161*, 813–820. https://doi.org/10.1016/j.procs.2019.11.187

Riccosan, Saputra, K. E., Pratama, G. D., & Chowanda, A. (2022). Emotion dataset from Indonesian public opinion. *Data in Brief*, *43*. https://doi.org/10.1016/j.dib.2022.108465

Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*. https://doi.org/10.1016/j.jocs.2019.05.009

Shaver, P. R., Murdaya, U., & Fraley, R. C. (2001). Structure of the Indonesian Emotion Lexicon. *Asian Journal of Social Psychology*, *4*(3), 201–224. https://doi.org/10.1111/1467-839X.00086

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6), 1061–1086. https://doi.org/10.1037/0022-3514.52.6.1061

Social media in Indonesia - statistics & facts | Statista. (n.d.). Retrieved July 9, 2023, from https://www.statista.com/topics/8306/social-media-in-indonesia/#topicOverview

Tapsell, R. (2021). Social Media and Elections in Southeast Asia: The Emergence of Subversive, Underground Campaigning. *Asian Studies Review*, *45*(1), 117–134. https://doi.org/10.1080/10357823.2020.1841093

Vaswani, A., Brain, G., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*.

Weismueller, J., Harrigan, P., Coussement, K., & Tessitore, T. (2022). What makes people share political content on social media? The role of emotion, authority

and ideology. *Computers in Human Behavior, 129.* https://doi.org/10.1016/j.chb.2021.107150

Nusa Putra
UNIVERSITY
2002
AMOR DEUS, PARENTIUM, CONCERVIS