



***SENTIMENT ANALYSIS CLASSIFICATION OF POLITICAL PARTIES ON TWITTER USING GATED RECURRENT UNIT ALGORITHM AND NATURAL LANGUAGE PROCESSING***

**TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat  
untuk Memperoleh Gelar Sarjana Komputer pada  
Program Studi Sistem Informasi

Oleh:

**AHMAD RIZKY ANDRIAWAN**

**11950314477**



UIN SUSKA RIAU

UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU  
PEKANBARU**

**2024**

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**LEMBAR PERSETUJUAN**

***SENTIMENT ANALYSIS CLASSIFICATION OF POLITICAL  
PARTIES ON TWITTER USING GATED RECURRENT UNIT  
ALGORITHM AND NATURAL LANGUAGE PROCESSING***

**TUGAS AKHIR**

Oleh:

**AHMAD RIZKY ANDRIAWAN**

**11950314477**

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir  
di Pekanbaru, pada tanggal 15 Januari 2024

**Ketua Program Studi**



**Eki Saputra, S.Kom., M.Kom.**

**NIP. 198307162011011008**

**Pembimbing**



**Mustakim, ST., M.Kom.**

**NIK. 130511023**

**LEMBAR PENGESAHAN**

***SENTIMENT ANALYSIS CLASSIFICATION OF POLITICAL PARTIES ON TWITTER USING GATED RECURRENT UNIT ALGORITHM AND NATURAL LANGUAGE PROCESSING***

**TUGAS AKHIR**

Oleh:

**AHMAD RIZKY ANDRIAWAN**

**11950314477**

Telah dipertahankan di depan sidang dewan penguji sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau di Pekanbaru, pada tanggal 08 Januari 2024

Pekanbaru, 08 Januari 2024

Mengesahkan,

Ketua Program Studi



**Eki Saputra, S.Kom., M.Kom.**

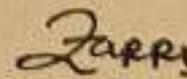
**NIP. 198307162011011008**



**Dr. Hartono, M.Pd.**  
**NIP. 196403011992031003**

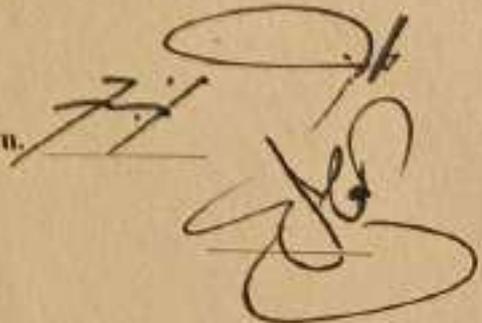
**DEWAN PENGUJI:**

Ketua : Zarnelly, S.Kom., M.Sc.



Sekretaris : Mustakim, ST., M.Kom.

Anggota 1 : Dr. Rice Novita, S.Kom., M.Kom.



Anggota 2 : M. Afdal, ST., M.Kom.

## LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum, dengan ketentuan bahwa hak cipta ada pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan atas izin peneliti dan harus dilakukan mengikuti kaedah dan kebiasaan ilmiah serta menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin tertulis dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan dapat meminjamkan Tugas Akhir ini untuk anggotanya dengan mengisi nama, tanda peminjaman dan tanggal pinjam pada *form* peminjaman.

### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



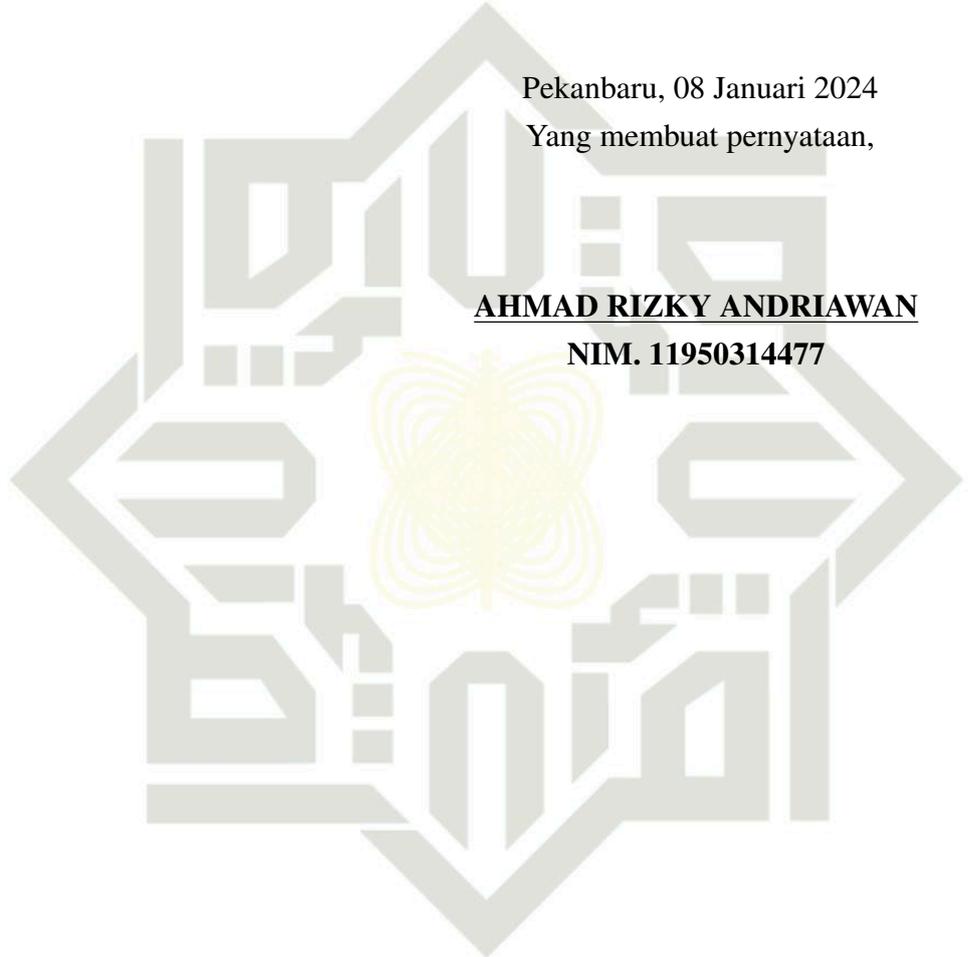
UIN SUSKA RIAU

## LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis dicantumkan dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 08 Januari 2024  
Yang membuat pernyataan,

**AHMAD RIZKY ANDRIAWAN**  
**NIM. 11950314477**



UIN SUSKA RIAU

### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LEMBAR PERSEMBAHAN



*Assalamu'alaikum Warahmatullahi Wabarakatuh.*

*Alhamdulillah Rabbil'Alamin*, puji dan syukur atas kehadiran Allah *Subhanahu Wa Ta'ala* sebagai bentuk rasa syukur atas segala nikmat yang telah diberikan tanpa ada kekurangan sedikitpun. Solawat beserta salam tidak lupa pula kita ucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan "*Allahumma Sholli'ala Muhammad Wa'ala Ali Muhammad*". Semoga kita semua senantiasa mendapat *syafa'at*-Nya di dunia dan akhirat kelak. Kupersembahkan karya kecil ini sebagai salah satu hadiah istimewa untuk orang tua tercinta sebagai rasa terima kasih dan hormat kepada keduanya.

Penulis mempersembahkan Tugas Akhir ini kepada orang yang sangat berarti dalam kehidupan penulis sebagai bentuk bakti penulis. Bapak dan ibu terkasih, terima kasih atas semua perjuangan, *do'a*, bimbingan, kasih sayang, dan dukungan yang telah diberikan kepada penulis dalam keadaan tersulit sekalipun. Semoga Allah *Subhanahu Wa Ta'ala* senantiasa selalu memberikan kesehatan dan perlindungan untuk bapak dan ibu, *Aamiin Ya Rabbal Alamiin*.

Kemudian penulis mengucapkan terima kasih kepada Bapak dan Ibu Dosen Program Studi Sistem Informasi yang telah mewariskan ilmu yang bermanfaat dan arahan kepada penulis untuk menyelesaikan pendidikan di Program Studi Sistem Informasi ini. Terima kasih juga kepada saudara-saudara penulis dan teman-teman seperjuangan yang selalu memberikan motivasi dan semangat kepada penulis hingga penulis dapat menyelesaikan Tugas Akhir ini. Semoga kita selalu dalam lindungan Allah *Subhanahu Wa Ta'ala*, *Aamiin*.

*Wassalamu'alaikum Warahmatullahi Wabarakaatuh.*

## KATA PENGANTAR

*Assalamu'alaikum Warahmatullahi Wabarakatuh.*

*Alhamdulillah Rabbil'Alamin*, puji dan syukur kehadiran Allah *Subhanahu Wa Ta'ala*, karena dengan Rahmat dan Karunia-Nya sehingga penulis dapat menyusun dan menyelesaikan Laporan Tugas Akhir ini yang berjudul "*Sentiment Analysis Classification of Political Parties on Twitter Using Gated Recurrent Unit Algorithm and Natural Language Processing*". Solawat beserta salam tidak lupa diucapkan kepada Nabi Muhammad *Shallallahu 'Alaihi Wa Sallam* dengan mengucapkan *Allahumma Sholli'ala Sayyidina Muhammad Wa'ala Ali Sayyidina Muhammad*.

Penulisan dan penyusunan Laporan Tugas Akhir ini tidak terlepas dengan adanya bantuan dari berbagai pihak, baik yang berupa materi maupun berupa motivasi. Untuk itu pada kesempatan ini penulis mengucapkan banyak terima kasih kepada:

1. Bapak Prof. Dr. Hairunas, M.Ag sebagai Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Hartono, M.Pd sebagai Dekan Fakultas Sains dan Teknologi.
3. Bapak Eki Saputra, S.Kom., M.Kom sebagai Ketua Program Studi Sistem Informasi.
4. Ibu Siti Monalisa, ST., M.Kom sebagai Sekretaris Program Studi Sistem Informasi.
5. Bapak Tengku Khairil Ahsyar, S.Kom., M.Kom sebagai Kepala Laboratorium Program Studi Sistem Informasi.
6. Bapak Inggih Permana, ST., M.Kom sebagai Dosen Pembimbing Akademik yang telah memberikan arahan, nasihat, dan motivasi kepada penulis selama perkuliahan.
7. Bapak Mustakim, ST., M.Kom sebagai Dosen Pembimbing Tugas Akhir yang telah memberikan arahan, masukan, dan nasihat dalam penyelesaian Tugas Akhir ini.
8. Ibu Zarnelly, S.Kom., M.Sc sebagai Ketua Sidang Tugas Akhir yang telah memberikan arahan, masukan, dan nasihat dalam penyelesaian Tugas Akhir ini.
9. Ibu Dr. Rice Novita, S.Kom., M.Kom sebagai Penguji I Tugas Akhir yang telah memberikan arahan, masukan, dan nasihat dalam penyelesaian Tugas Akhir ini.
10. Bapak M. Afdal, ST., M.Kom sebagai Penguji II Tugas Akhir yang telah

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

memberikan arahan, masukan, dan nasihat dalam penyelesaian Tugas Akhir ini.

1. Para Dosen Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau yang memberikan ilmu yang bermanfaat serta memberikan semangat kepada penulis. Semoga ilmu yang diberikan dapat penulis amalkan dalam kehidupan sehari-hari.
2. Yang istimewa terima kasih kepada kedua orang tua yaitu ayah terhebat dan ibu tercinta sebagai motivator utama dalam penyelesaian Tugas Akhir ini yang telah memberikan dukungan, *do'a*, perhatian, kasih sayang, dan semangat yang tiada hentinya. Semoga Allah selalu melindungi dan memberikan kesehatan kepada ayah dan ibu, *Aamiin Ya Rabbal'alamiin*.
3. Terima kasih kepada kakak dan adik tersayang sebagai saudara kandung penulis, dan keluarga besar yang telah memberikan semangat, perhatian, motivasi, *do'a*, dan dukungan sehingga penulis mampu menyelesaikan Tugas Akhir ini.
14. Kepada teman-teman *Puzzle Research Data Technology* terdiri dari Akhas, Bobby, Adyah, Sania, Umairah, Hilda, Rahweni, Ratna, Ainun, dan Frendi serta teman-teman *Predatech* lainnya yang memberi inspirasi dan motivasi kepada penulis.
15. Kepada sahabat seperjuangan Kelas Unggul 2019 yang telah berjuang bersama dalam menuntut ilmu.

Semoga segala *do'a* dan motivasi yang telah diberikan selama ini menjadi amal kebajikan dan mendapatkan pahala terbaik dari Allah *Subhanahu Wa Ta'ala*. Penulis menyadari bahwa penulisan Tugas Akhir ini masih terdapat kekurangan, untuk itu kritik dan saran yang bersifat membangun dari semua pihak untuk kesempurnaan Tugas Akhir ini. Penulis berharap Laporan Tugas Akhir ini bermanfaat bagi pembaca dan kita semua khususnya pada bidang Sistem Informasi. Akhir kata penulis mengucapkan terima kasih.

*Wassalamu'alaikum Warahmatullahi Wabarakatuh.*

Pekanbaru, 16 Januari 2024

Penulis,

**AHMAD RIZKY ANDRIAWAN**

**NIM. 11950314477**

: 109/UMA/JITE/XI/2023

Medan, 24 November 2023

: Letter of Acceptance

Mr. Ms. **Ahmad Rizky Andriawan**

Assalamu'alaikum Wr. Wb

I would like to express our sincere gratitude for your participation in submitting an article to the Journal of Informatics and Telecommunication Engineering (JITE). We hereby inform you that the article listed below:

<b>Paper</b>	: <b>Sentiment Analysis Classification Of Political Parties On Twitter Using Gated Recurrent Unit Algorithm And Natural Language Processing</b>
<b>Authors</b>	: <b>Ahmad Rizky Andriawan*, Mustakim &amp; Rice Novita</b>

Based on the review results, the article you submitted has been ACCEPTED to be published in JITE Journal Volume 7 Number 2, January 2024, ISSN: 2549-6247 (Print) ISSN: 2549-6255 (Online).

I would like to thank you for your attention and cooperation.

Assalamu'alaikum Wr. Wb.

Best Regards,



**Muhathir, ST., M.Kom**

Chief Editor

UIN SUSKA RIAU

2. Di larang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak cipta milik UIN Suska Riau State Islamic University of Sultan Syarif Kasim Riau

Lampiran Surat :

Nomor : Nomor 25/2021

Tanggal : 10 September 2021

### SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Ahmad Rizky Andriawan

NIM : 11950314477

Tempat/ Tgl. Lahir : TTL: Bukit Lipai, 2 Juni 2001

Fakultas/Pascasarjana : Sains dan Teknologi

Prodi : Sistem Informasi

Judul Disertasi/Thesis/Skripsi/Karya Ilmiah lainnya\*:

Scntiment Analysis Classification Of Political Parties On Twitter Using Gated Recurrent Unit

Algorithm And Natural Language Processing

---

---

---

---

---

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan Disertasi/Tesis/Skripsi/Karya Ilmiah lainnya\* dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu Disertasi/Tesis/Skripsi/Karya Ilmiah lainnya\* saya ini, saya nyatakan bebas dari plagiat.
4. Apa bila dikemudian hari terbukti terdapat plagiat dalam penulisan Disertasi/Tesis/Skripsi/(Karya Ilmiah lainnya)\* saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikian Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga.

Pekanbaru, 15 Januari 2024

, membuat pernyataan



Ahmad Rizky Andriawan

NIM : 11950314477

Hak Cipta Diindungi Undang-Undang  
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:  
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tirjauan suatu masalah.  
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.  
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

# JITE (Journal of Informatics and Telecommunication Engineering)

Available online <http://ojs.uma.ac.id/index.php/jite> DOI: 10.31289/jite.vxiv.xxx

Received: dd-mm-yyyy

Accepted: dd-mm-yyyy

Published: dd-mm-yyyy

## Sentiment Analysis Classification Of Political Parties On Twitter Using Gated Recurrent Unit Algorithm And Natural Language Processing

Ahmad Rizky Andriawan1)\*, Mustakim2) & Rice Novita3)

1,2,3) Department of Information System, Faculty of Science and Technology, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

\*Corresponding Email: 11950314477@students.uin-suska.ac.id

### Abstrak

Pemilihan umum tidak lepas kaitannya dengan isu partai politik. Isu tersebut dapat berupa survei hingga sentimen. Hasil dari survei saat ini perlu dilakukan validasi mendalam terkait dengan kebenarannya. Analisis sentimen bertujuan untuk memvalidasi akan kebenaran dari lembaga survei tersebut. Terdapat 5 partai politik yang digunakan sebagai dataset pada penelitian ini yaitu Partai Demokrasi Indonesia Perjuangan (PDIP), Partai Gerakan Indonesia Raya (Gerindra), Partai Golongan Karya (Golkar), Partai Kebangkitan Bangsa (PKB), dan Partai Nasional Demokrat (Nasdem). Algoritma Gated Recurrent Unit (GRU) diimplementasikan pada penelitian ini sebagai percobaan dalam penghitungan data. Berdasarkan hasil dari penghitungan algoritma GRU dalam menghitung sentimen pada partai politik menghasilkan data tertinggi pada akurasi 56,50%, presisi 72,76%, dan recall 100%.

**Kata Kunci:** Natural Language Processing, Gated Recurrent Unit, Partai Politik, Analisa Sentimen, Twitter

### Abstract

General elections cannot be separated from the issue of political parties. The issue can be in the form of surveys to sentiment. The results of the current survey need to be done in-depth validation related to the truth. Sentiment analysis aims to validate the truth of the survey institution. There are 5 political parties used as datasets in this study, namely Partai Demokrasi Indonesia Perjuangan Party (PDIP), Gerakan Indonesia Raya Party (Gerindra), Golongan Karya Party (Golkar), Partai Kebangkitan Bangsa Party (PKB), and Nasional Demokrat Party (Nasdem). The Gated Recurrent Unit (GRU) algorithm is implemented in this research as an experiment in data calculation. Based on the results of the GRU algorithm calculation in calculating sentiment on political parties, it produces the highest data at 56.50% accuracy, 72.76% precision, and 100% recall.

**Keywords:** Natural Language Processing, Gated Recurrent Unit, Political Parties, Sentiment Analysis, Twitter

### How to Cite:

## I. INTRODUCTION

Reporting from the official website of the General Elections Commission (KPU), [kpu.go.id](http://kpu.go.id) (30/12/2022), there are 24 parties that will participate in the 2024 elections. Based on data from [pemilu2019.kpu.go.id](http://pemilu2019.kpu.go.id), there are parties that occupy the top 5 ranks, namely Partai Demokrasi Indonesia Perjuangan Party (PDIP), Gerakan Indonesia Raya Party (Gerindra), Golongan Karya Party (Golkar), Partai Kebangkitan Bangsa Party (PKB), and Nasional Demokrat Party (Nasdem). Thus, the party will take part as an election participant in the 2024 general election.

The credibility of survey institutions in Indonesia needs to be validated. Reporting from [kompas.id](http://kompas.id) (28/03/2023), according to Titi Anggraini as a lecturer in election law at the Faculty of Law, University of Indonesia, said that survey results can be signaled as manipulation carried out by fake survey providers which do not reflect actual conditions. The results of a survey are also influenced by different methodologies, samples, and data collection areas and do not represent the Indonesian people as a whole. Thus, the results of the survey must also be different. Meanwhile, sentiment analysis on twitter social media is considered more representative because it is based on the perceptions of the community directly. So, this research will

prove the success of survey institutions and as a comparison to sentiment analysis, which is a direct national opinion through social media intermediaries.

Twitter is a social media platform that is popularly used among internet users (Phan et al., 2020) (Curlin et al., 2019) (Giovani et al., 2020). Extensive public data sources span many studies (Karami et al., 2020), mostly used for text mining research, one of them is related to politics. Research using Twitter as a dataset has been conducted by (Nurul et al., 2019) in analyzing public sentiment towards Indonesian presidential candidates in 2019. In terms of social politics, Twitter is widely used as a medium for expressing opinions and conducting campaigns and can be used as a place by the public to provide opinions on related parties (Adiati et al., 2019). The openness of users in providing opinions, opinions, suggestions, and criticisms on Twitter provides an opportunity for the author to explore data related to political parties used as research datasets.

In analyzing sentiment, the use of the Gated Recurrent Unit (GRU) algorithm has been done by (Raza et al., 2021) in comparing the accuracy of the Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and GRU algorithms, the results show that GRU is the best among the three. In research (Sachin et al., 2020) in comparing the performance of the GRU and LSTM algorithms, resulting in the GRU algorithm performing better than LSTM. The use of the GRU algorithm has also been implemented in analyzing the sentiment of the presidential election in Indonesia in 2019, where the GRU algorithm produces an accuracy rate of 84.50% (Hidayatullah et al., 2021). With a fairly high accuracy in previous studies, the GRU algorithm is feasible to use in this study in calculating data accuracy.

Natural Language Processing (NLP) is used to process data and corpus so that machines can understand it (Munasatya & Novianto, 2020) (Otter et al., 2019). The use of NLP has been used by (Pasek et al., 2022) in analyzing sentiment towards the leader figure, Anies Baswedan, on twitter social media. Research related to the 2024 Election using NLP has also been carried out by (Vindua & Zailani, 2023) using the keyword "Pemilu 2024". With the success and the amount of literature using NLP, it is the reason for the author to use NLP as a research method.

With the momentum of elections in Indonesia that will be held in 2024, of course, political parties play a role in the implementation of elections. In this case, the party selects prospective leaders who are eligible to hold positions. During this period, political parties also received more attention from the public, especially on twitter social media. People expressed their criticisms, suggestions, and opinions related to the party either in a positive, negative, or neutral context. Therefore, it is necessary to classify and analyze political parties in Indonesia, especially the 5 parties with the highest rankings as a result of the 2019 elections, namely the PDI-Perjuangan Party (PDIP), Gerindra Party, Golkar Party, PKB Party, and Nasdem Party.

## II. RESEARCH METHOD

The stages in this research start from data retrieval (Crawling), Preprocessing, to the analysis and results stage. These stages can be seen in Fig. 1.

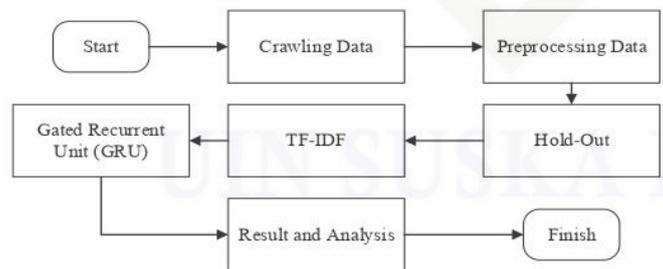


Fig. 1. Research Methodology

In this research, the first stage is to collect data which is commonly referred to as crawling data (Normawati & Prayogi, 2021). Data that has been crawled is preprocessed by cleaning, tokenizing, filtering, and stemming (Deolika & Taufiq Luthfi, 2019). Next, perform data transformation, namely the TF-IDF stage. Clean data is compared with three data divisions, namely testing data of 10%, 20%, and 30% respectively, and training data of 90%, 80%, and 70% respectively. The algorithm used in this research is Gated Recurrent Unit (GRU) in calculating the performance of accuracy, precision, and recall. The last stage is results and



analysis, which describes the results of data calculation from preprocessing to the GRU algorithm calculation stage.

### A. Data Crawling

*Data crawling* is the process of collecting data from Twitter (Normawati & Prayogi, 2021). Based on the case that will be used in this research are 5 political parties in Indonesia, namely the PDI-Perjuangan Party (PDIP), Gerindra Party, Golkar Party, PKB Party, and Nasdem Party.

### B. Hold Out Validation

Hold out is currently considered one of the most reliable data sharing techniques for prediction model accuracy (Lee et al., 2018) (Rahmadeyan et al., 2023). Hold out is also used for data division into two data sets: training data and testing validation data (Rafał, 2022) (Ramaswami et al., 2019). The training data is utilized to train a classification model. Meanwhile, the test data is utilized to evaluate the resulting model (Gurning et al., 2022).

### C. TF-IDF

The TF-IDF method is used to calculate the weight value of the term and the sentence with the highest value to be displayed in the summary (Andriani & Tanzil Furqon, 2019). TF (Term Frequency) is the frequency of appearance of a term in a document. The calculation of TF value can be done with Equation (1).

$$\begin{cases} 1 + \log_{10}(f_{t,d}), f_{t,d} > 0 \\ 0, f_{t,d} = 0 \end{cases} \quad (1)$$

Meanwhile, IDF (Inverse Document Frequency) is a calculation of how a term is distributed in a document collection. The calculation of the IDF value can be done with Equation (2).

$$IDF = \log\left(\frac{N}{df}\right) \quad (2)$$

### D. Gated Recurrent Unit

GRU combines forget gate and input gate into one update gate. GRU also incorporates cell states and hidden states, and some other changes. The final model is simpler than the standard LSTM unit (Liu et al., 2020). GRU is a type of recurrent artificial neural network, which is used in sequence modeling problems and is particularly suitable for NLP. GRU capture networks involve fewer parameters so they are faster to run (Shrivastava & Kumar, 2020). GRU is a relatively new development made by Cho et al. Similar to the LSTM function, GRU consists of a gating unit that controls the movement of data and information within the unit. GRU further computes two important gates called update and reset gates that modulate the movement of information entering each hidden unit (Sachin et al., 2020).

$$\text{Update gate: } z_t = \sigma_g(W x^{(z)}_t + U h^{(z)}_{t-1} + b^{(z)}) \quad (3)$$

$$\text{Reset gate: } r_t = \sigma_g(W x^{(r)}_t + U h^{(r)}_{t-1} + b^{(r)}) \quad (4)$$

$$\text{Candidate state: } \hat{h}_t = \tanh(W x^{(\hat{h})}_t + U^{(\hat{h})}(r_t * h_{t-1})) \quad (5)$$

$$\text{Final Output: } h_t = z_t * h_{t-1} + (1 - z_t) * \hat{h}_t \quad (6)$$

### E. Performance Measurement

Performance measurement aims to determine how well the algorithm is used. The measurement uses accuracy, precision, and recall parameters (Izzah et al., 2020) (Muktafin et al., 2020).

Accuracy is the measurement of correctly classified samples compared to the total number of samples.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (7)$$

Precision is the proportion of correctly classified positive samples to the total number of samples predicted to be positive.

$$\text{Precision} = \frac{TP}{FP+TP} \quad (8)$$

Recall is a measurement of the positives ratio that are correctly classified as positive.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (9)$$



III. RESULT AND DISCUSSION

**A. Data Crawling**

Data is obtained by data crawling which is a data collection technique from twitter. This research uses a total of 10,000 data from twitter social media with a crawling period from January 01, 2023 to September 20, 2023. Crawled data samples can be seen in Table 1.

Table 1. Result of Crawling Data

	Tweet
0	Bu Puan Menunjukkan bahwa humor adalah bahasa ...
1	Beginilah. Kalau Partai PDI Perjuangan Yang Me...
2	gimane coba, katenye mau mengusung perubahan, ...
3	PDI Perjuangan lebih banyak ka, satu2 nya part...
...	...
1999	@PDI_Perjuangan Membawa gerbong isinya hantu.

**B. Data Preprocessing**

Preprocessing is a stage carried out in text processing to form a good data structure. At this stage the cleaned data is transformed into TF-IDF form in Table 2.

Table 2. TF-IDF Data

begitu	begitulah	bego	beijing	bekerja	bela
0	0	0	0	0,194	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0,398	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0,241
0	0	0	0	0	0,188
0	0	0	0	0	0,289
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0,245	0
0	0	0	0	0	0
0	0	0,643	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0,146	0
0	0	0	0	0	0
0	0	0	0	0	0
0,193	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0,290	0
0	0	0	0	0	0
0	0	0	0	0	0

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

0	0	0	0	0,241	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0,387	0
0	0	0	0	0,189	0
0	0	0	0	0,373	0
0	0	0	0	0	0
0	0	0	0	0,580	0

### C. Algorithm Calculations

In this section there are several stages, namely data classification based on positive, negative, and neutral sentiments with a total of 2000 data for each party that has gone through the previous preprocessing stage. Then the data is calculated for accuracy using the GRU algorithm.

#### 1. Data Classification

Data that has gone through the preprocessing process is classified based on its sentiment with positive, negative, and neutral variables. The amount of data that has been classified can be seen in Table 3.

Table 3. Data Classification

Party	Positive	Negative	Neutral
PDI-P	746	898	356
Gerindra	1087	643	270
Golkar	1036	396	568
Nasdem	721	669	610
PKB	991	443	566

Based on the amount of data in Table 3, the most positive data on the Gerindra party is 1087 data, the most negative data on the PDI-P party is 898 data, and neutral data on the Nasdem party is 610 data.

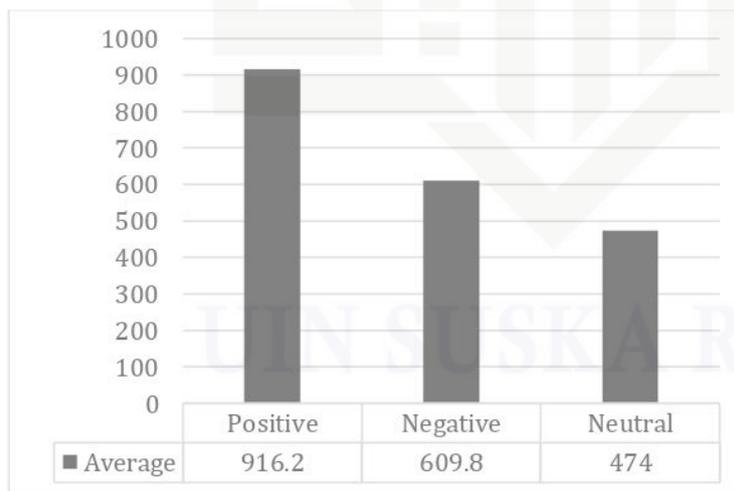


Fig. 2. Average Amount of Data in Each Sentiment

Based on Fig. 2, of all political parties there is more data from positive sentiment with an average of 916.2, then negative with an average of 609.8, and neutral at 474.

#### 2. GRU Algorithm Calculations

After the data is classified with positive, negative, and neutral variables, the next process is to calculate performance, namely accuracy, precision, and recall using the GRU algorithm.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

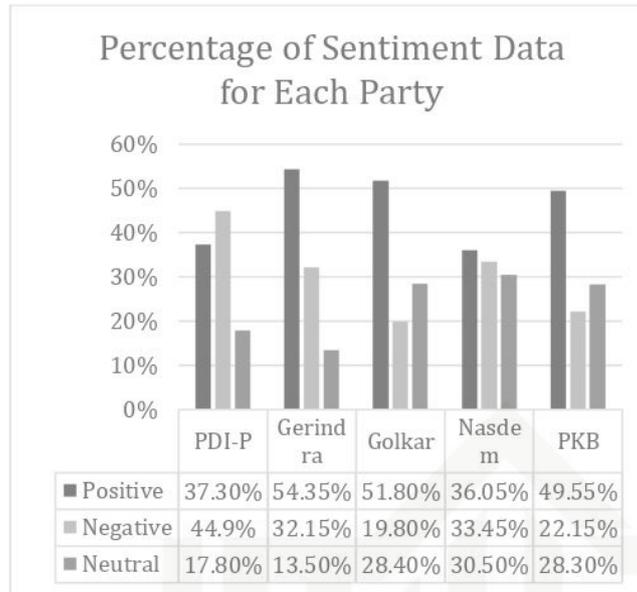


Fig. 3. Percentage of Sentiment Data Political Parties in Each Data

Based on the amount of data in Fig. 3, the most positive data on the Gerindra party is 54.35%, the most negative data on the PDI-P party is 44.9%, and neutral on the Nasdem party is 30.50%.

Table 4. Performance Measurements Comparison

Hold Out	Party	Accuracy	Precision	Recall
70:30	PDI-P	31,17%	71,07%	91,87%
	Gerindra	14,67%	48,83%	100%
	Golkar	55,17%	69,97%	76,04%
	Nasdem	28,33%	62,67%	100%
	PKB	27,33%	49,17%	100%
80:20	PDI-P	32,00%	72,76%	87,95%
	Gerindra	40,00%	64,89%	72,64%
	Golkar	56,50%	71,56%	80,41%
	Nasdem	29,25%	61,50%	100%
	PKB	26,50%	48,00%	100%
90:10	PDI-P	18,50%	66,50%	100%
	Gerindra	14,00%	46,50%	100%
	Golkar	29,50%	53,50%	100%
	Nasdem	30,00%	61,50%	100%
	PKB	27,50%	51,50%	100%

The implementation of the GRU algorithm in Table 4 produces accuracy that is not too high when compared to previous studies (Hidayatullah et al., 2021). As stated in Table 4, the highest accuracy in this study is on the Golkar party dataset of 56.50%. Meanwhile, the highest precision on the PDI-P dataset is 72.76%. Then, recall is 100% on the Nasdem and PKB datasets, and on the Gerindra dataset with 30% testing data.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya atau ini tanpa mencantumkan dan menyebutkan sumber.
  - a. Penguipatan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Penguipatan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

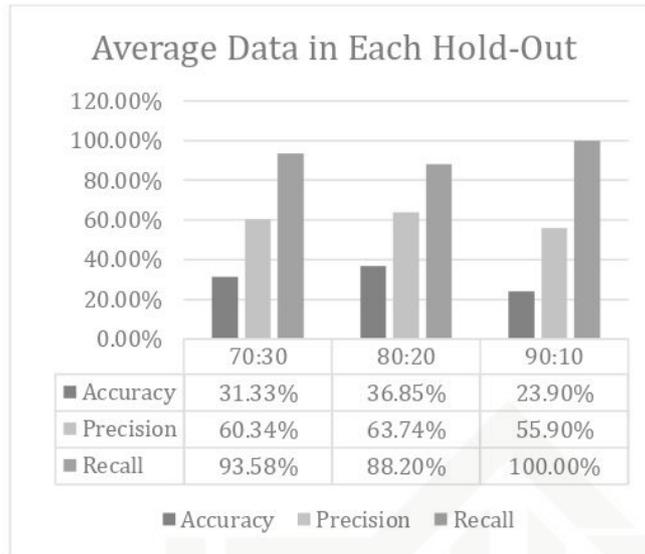


Fig. 4. Percentage of Average Data in Each Hold-Out

Based on Fig. 4, the average percentage with the highest accuracy on 20% testing data is 36.85%, while the highest average precision on 20% testing data is 63.74%, and the highest average recall on 10% testing data is 100%. So, overall data sharing using hold out shows the best results using 20% testing data.



Fig. 5. The Highest Positive Sentiment From The Gerindra Party

It can be seen in Fig. 5, that in the dataset with the highest positive sentiment, namely the Gerindra party data, there are 5 most common words that appear, namely "Gerindra", "prabowo", "ini", "mantap", and "semoga". It can be concluded that there are prayers, hopes, and support addressed to prabowo who is in the Gerindra party characterized by the appearance of the words "semoga" and "mantap".

the results show the highest positive sentiment dataset from the Gerindra Party. which, these results can be compared and become a picture with survey institutions related to the Gerindra Party in the future. the picture and comparison can be in the form of how high the level of positive and negative sentiment is. therefore, opening people's views regarding their choices in the 2024 elections.



Fig. 6. The Highest Negative Sentiment From The PDI-P Party



Based on Figure 6, in the dataset with the highest negative sentiment, namely the PDI-P party data, there are 4 most common words that appear, namely "tenggelamkan", "kadrun", "jangan", and "partai". It can be concluded that there is a negative word namely "kadrun" and negative invitation characterized by the words "jangan" and "tenggelamkan", which is related to PDI-P party.

#### D. Discussion

Although the lower accuracy and gain values indicate that GRU may underperform in this scenario, other factors should be considered. When compared to previous studies using the GRU algorithm by (Hidayatullah et al., 2021), this study obtained a fairly low accuracy. This can be caused by the data used in the study being less suitable for the GRU algorithm, the quality of the dataset, and the amount of data used. This can cause performance to be less than optimal. thus, the features used for classification must be truly independent.

#### IV. CONCLUSION

Based on testing a dataset of 5 political parties in this study, namely the PDI-Perjuangan Party (PDIP), Gerindra Party, Golkar Party, PKB Party, and Nasdem Party using the GRU algorithm, the highest accuracy was 56.50%, the highest precision was 72.76%, and recall was 100%. With these accuracy results, it can be concluded that this research using the GRU algorithm and political parties data in Indonesia produces poor results when viewed from the resulting accuracy. This can occur due to several things related to the algorithm and the data used.

Therefore, in future research, an in-depth study of the GRU algorithm is needed which is still relatively new and there is still a lack of related research. experiments with other algorithms can be carried out such as the Convolutional Neural Network algorithm, Recurrent Neural Network, or Long Short Term Memory in research related to sentiment analysis. then, in addition to using hold out cross validation, other cross validation methods can also be used such as K-Fold Cross Validation or Stratified K-Fold Cross Validation, in order to get an assessment of each method and to assess the performance and generalization ability of a model.

#### REFERENCE

- Adiati, A. R., Herdiani, A., Astuti, W., & Kom, M. (2019). *Analisis Sentimen Masyarakat Pada Media Sosial Twitter Terhadap Partai Politik Peserta Pemilu 2019 Menggunakan Naïve Bayes Classifier*.
- Andriani, D., & Tanzil Furqon, M. (2019). Peringkasan Teks Otomatis Pada Artikel Berita Hiburan Berbahasa Indonesia Menggunakan Metode Bm25. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 3(3), 2548–2964. <http://j-ptiik.ub.ac.id>
- Curlin, T., Jaković, B., & Miloloža, I. (2019). Twitter Usage In Tourism: Literature Review. *Business Systems Research*, 10(1), 102–119. <https://doi.org/10.2478/bsrj-2019-0008>
- Deolika, A., & Taufiq Luthfi, E. (2019). Analisis Pembobotan Kata Pada Klasifikasi Text Mining. *Jurnal Teknologi Informasi*, 3(2).
- Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*, 14(2), 115. <https://doi.org/10.33365/jti.v14i2.679>
- Gurning, U. R., Mustakim, Rizaldi, S. T., & Syukron, H. (2022). Comparison Of Naïve Bayes, C4.5 And K-Nearest Neighbor For Covid-19 Data Classification. *2022 International Symposium On Information Technology And Digital Innovation (Isitdi)*, 16–21. <https://doi.org/10.1109/Isitdi55734.2022.9944507>





Raza, M. R., Hussain, W., & Merigo, J. M. (2021). Cloud Sentiment Accuracy Comparison Using Rnn, Lstm And Gru. *Proceedings - 2021 Innovations In Intelligent Systems And Applications Conference, Asyu 2021*. <https://doi.org/10.1109/Asyu52992.2021.9599044>

Sachin, S., Tripathi, A., Mahajan, N., Aggarwal, S., & Nagrath, P. (2020). Sentiment Analysis Using Gated Recurrent Neural Networks. In *Sn Computer Science* (Vol. 1, Issue 2). Springer. <https://doi.org/10.1007/S42979-020-0076-Y>

Shrivastava, K., & Kumar, S. (2020). A Sentiment Analysis System For The Hindi Language By Integrating Gated Recurrent Unit With Genetic Algorithm. *International Arab Journal Of Information Technology*, 17(6), 954–964. <https://doi.org/10.34028/Iajit/17/6/14>

Vindua, R., & Zailani, A. U. (2023). Analisis Sentimen Pemilu Indonesia Tahun 2024 Dari Media Sosial Twitter Menggunakan Python. *Jurikom (Jurnal Riset Komputer)*, 10(2), 479. <https://doi.org/10.30865/Jurikom.V10i2.5945>



UIN SUSKA RIAU

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Peningkatan mutu dan kualitas penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
 b. Penguatan hak merdeka kepentingan yang wajar UIN Suska Riau.

Dilarang mengizinkan seluruh atau sebagian dari seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:  
 a. Peningkatan mutu dan kualitas penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.  
 b. Penguatan hak merdeka kepentingan yang wajar UIN Suska Riau.

## LAMPIRAN A

### NOTIFIKASI PAPER

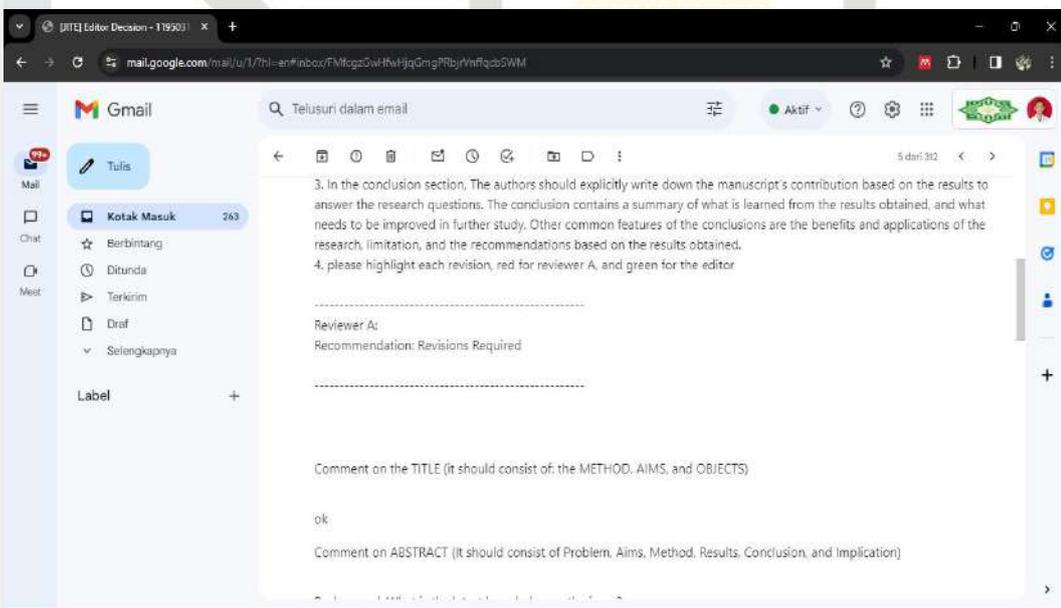
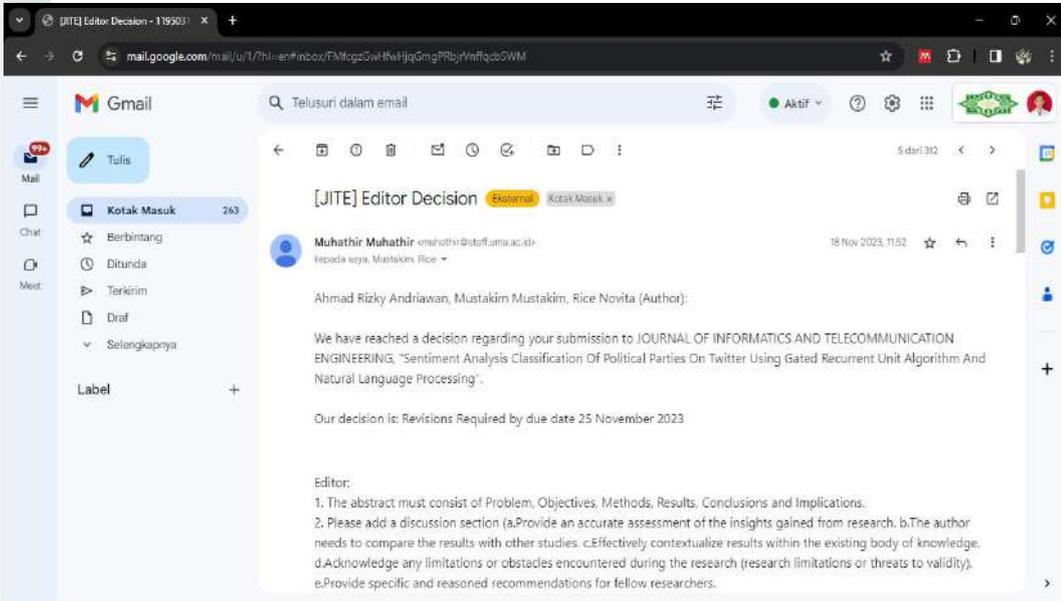
Hak Cipta Diindungi Undang-Undang

Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

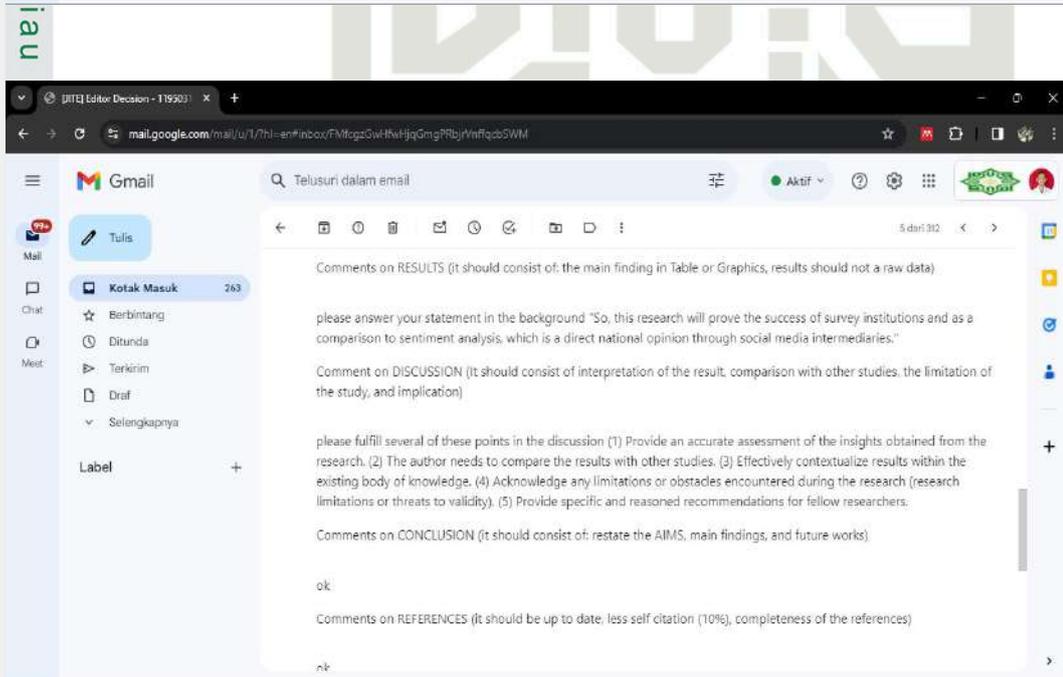
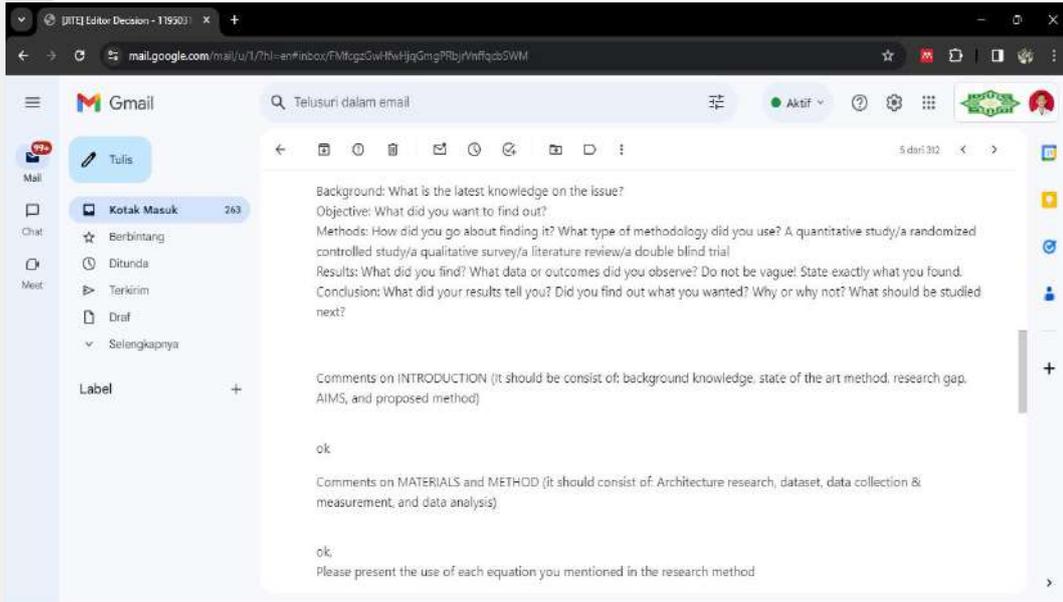
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

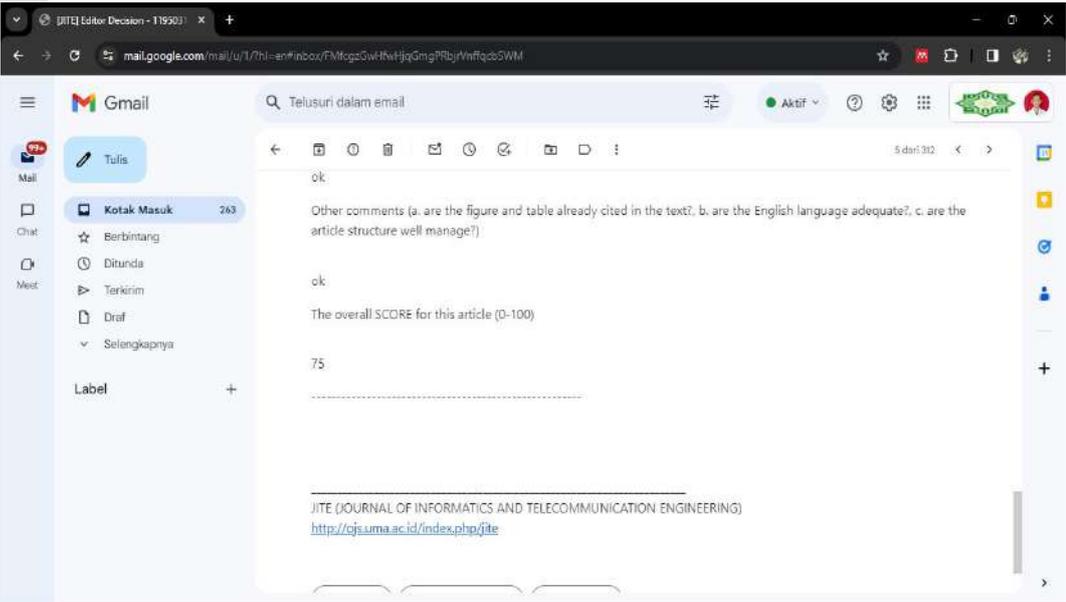


Hak Cipta Diindungi Undang-Undang

### Hak Cipta Diindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.





iau

State Islamic University of Sultan Syarif Kasim Riau

### Hak Cipta Diindungi Undang-Undang

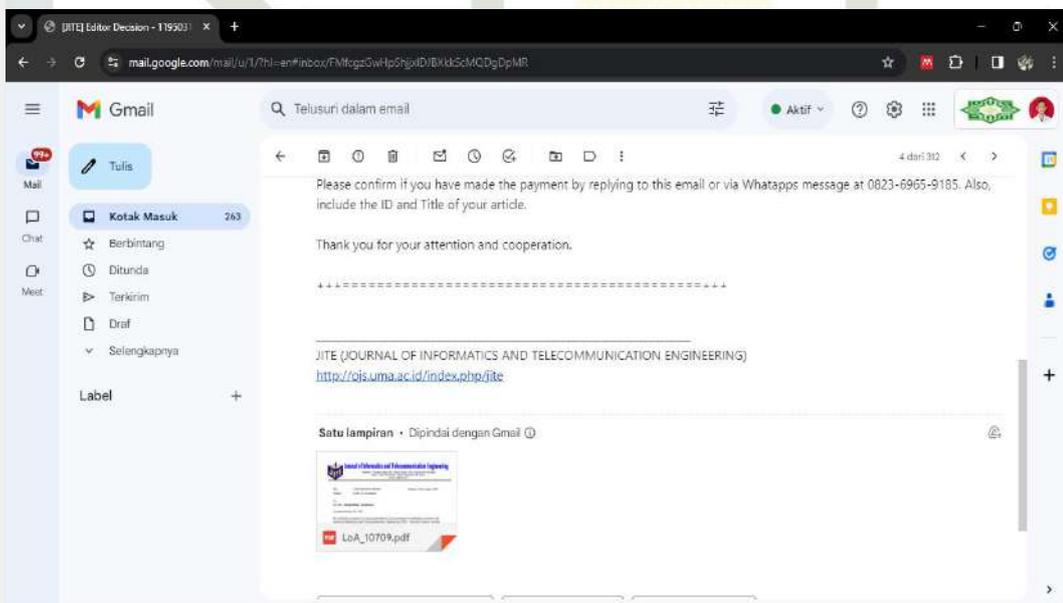
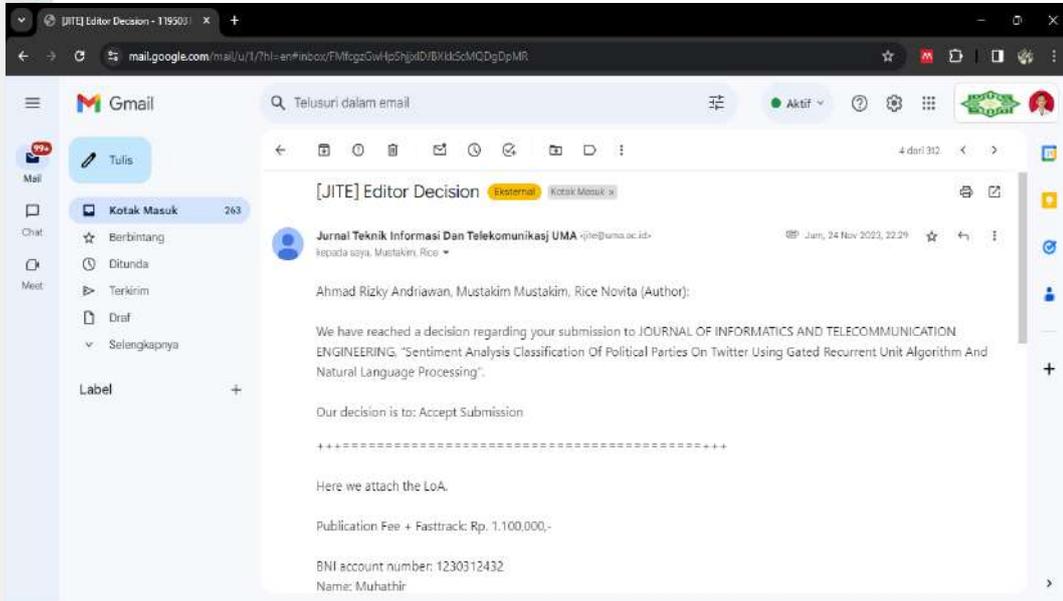
1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## LAMPIRAN B ACCEPTED PAPER

Hak C

### Hak Cipta Diindungi Undang-Undang

1. Diarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Diarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



of Sultan Syarif Kasim Riau

## SURAT PERNYATAAN PUBLIKASI JURNAL

Saya yang bertanda tangan dibawah ini:

Nama : Ahmad Rizky Andriawan

NIM : 11950314477

Jurusan : Sistem Informasi

Judul Tugas Akhir : Sentiment Analysis Classification of Political Partics on Twitter Using Gated Recurrent Unit Algorithm and Natural Language Processing

Dengan ini menyatakan bahwa akan melengkapi seluruh kelengkapan administrasi Tugas Akhir Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau berupa bukti Letter of Acceptance (LoA) dan email pengiriman LoA tersebut. Demikian surat pernyataan ini dibuat untuk dipergunakan sebagaimana mestinya.

Pekanbaru, 14 Januari 2024

Yang membuat pernyataan



Ahmad Rizky Andriawan

NIM. 11950314477



## DAFTAR RIWAYAT HIDUP

Penulis lahir di Bukit Lipai pada tanggal 02 Juni 2001. Penulis anak dari Bapak Kotib dan Ibu Rif Ngah yang bernama Ahmad Rizky Andriawan. Penulis bertempat tinggal di Gang Satria, Jl. Garuda Sakti, Simpang Baru, Tampan, Pekanbaru, Riau. Penulis menempuh pendidikan yang diawali dari pendidikan di TK Pertiwi Desa Bukit Lipai. Selanjutnya melanjutkan pendidikan di SDN 009 Bukit Lipai pada tahun 2007-2013. Setelah menempuh pendidikan Sekolah Dasar, penulis melanjutkan pendidikan di MTS Al-Multazam Desa Kerubung Jaya tahun 2013-2016, dan dilanjutkan di SMK N 1 Rengat tahun 2016-2019 dengan Jurusan Teknik Komputer dan Jaringan (TKJ). Setelah menyelesaikan pendidikan dibangku sekolah, penulis melanjutkan pendidikan Strata Satu (S1) di Universitas Islam Negeri Sultan Syarif Kasim Riau, tepatnya Fakultas Sains dan Teknologi Jurusan Sistem Informasi tahun 2019. Selama masa perkuliahan, penulis bergabung dalam organisasi kampus yaitu *Puzzle Research Data Technology* (Predatech). Dalam organisasi Predatech, penulis banyak mendapatkan pengalaman sebagai Anggota hingga Wakil Koordinator Mahasiswa 2022/2023. Pada penelitian Tugas Akhir, penulis mengambil topik data mining dengan judul "*Sentiment Analysis Classification of Political Parties on Twitter Using Gated Recurrent Unit Algorithm and Natural Language Processing*". Untuk menjalin komunikasi yang baik dengan penulis baik dari dalam kampus maupun luar kampus dapat menghubungi kontak melalui email [11950314477@students.uin-suska.ac.id](mailto:11950314477@students.uin-suska.ac.id).

### Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.