

Analisis Sentimen Komentar YouTube pada Video Clash of Champions di Channel Ruanguru Menggunakan Algoritma BERT

Laporan Tugas Akhir

Diajukan Untuk Memenuhi

Persyaratan Guna Meraih Gelar Sarjana Informatika

Universitas Muhammadiyah Malang



Muhammad Galieh Mardiono

202010370311442

Bidang Minat

Data Science

**PROGRAM STUDI INFORMATIKA FAKULTAS
TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2024**

LEMBAR PERSETUJUAN

**Analisis Sentimen Komentar YouTube pada Video Clash of Champions di
Channel Ruang Guru Menggunakan Algoritma BERT**

TUGAS AKHIR

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang**

Menyetujui,
Malang, 6 Maret 2025

Dosen Pembimbing 1



Christian Sri Kusuma Aditya S.Kom., M.Kom

NIP. 180327021991PNS.

Dosen Pembimbing 2



Didih Rizki Chandranegara S.kom.,

M.Kom

NIP. 180302101992PNS.

LEMBAR PENGESAHAN

Analisis Sentimen Komentar YouTube pada Video Clash of Champions di Channel Ruang Guru Menggunakan Algoritma BERT

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

MUHAMMAD GALIEH MARDIONO

202010370311442

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis pengujian
pada tanggal 6 Maret 2025

Menyetujui,

Dosen Penguji 1



Ir. Galih Wasis Wicaksono S.kom.

M.Cs.

NIP. 10814100541PNS.

Dosen Penguji 2



Vinna Rahmayanti S S.Si., M.Si

NIP. 180306071990PNS.



Mengetahui,
Ketua Jurusan Informatika

Ir. Galih Wasis Wicaksono S.kom. M.Cs.

NIP. 10814100541PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : **Muhammad Galieh Mardiono**

NIM : **202020370311442**

FAK/JUR. : **TEKNIK/INFORMATIKA**

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul “ **Analisis Sentimen Komentar YouTube Pada Video Clash of Champions di Channel Ruanguru Menggunakan Algoritma BERT** “ beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik Sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini, maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,

Dosen Pembimbing 1



(Christian Sri Kusuma, S.Kom,M.Kom)

Malang,22 Janurari 2025

Yang Membuat Pernyataan



(Muhammad Galieh Mardiono)

ABSTRAK

Analisis sentimen merupakan teknik yang bertujuan untuk memahami opini masyarakat melalui teks, termasuk komentar yang terdapat pada video YouTube. Komentar pengguna pada platform tersebut sering kali menjadi representasi beragam pandangan publik terhadap konten tertentu. Penelitian ini berfokus pada analisis sentimen komentar YouTube pada video *Clash of Champions* di *Channel* Ruangguru dengan menerapkan model *Bidirectional Encoder Representations from Transformers* (BERT). Tujuan dari penelitian ini adalah menganalisis sentimen komentar menggunakan model BERT, yang dinilai memiliki keunggulan dalam memahami bahasa alami secara mendalam. Dataset penelitian terdiri dari komentar YouTube yang telah melalui tahapan *pre-processing*, pelabelan manual, serta augmentasi data guna memastikan kualitas analisis yang optimal. Dalam implementasinya, model BERT diadaptasi menggunakan library *indobenchmark/indobert-base-p1* dan dievaluasi berdasarkan metrik akurasi, presisi, *recall*, serta skor F1. Hasil penelitian menunjukkan bahwa model ini mampu mencapai akurasi hingga 96% dalam mengklasifikasikan sentimen ke dalam empat kategori: positif, negatif, netral, dan kategori lain (*others*). Temuan dari studi ini dapat membantu Ruangguru memahami opini publik secara lebih mendalam sekaligus memberikan masukan strategis untuk meningkatkan kualitas konten yang disajikan.

Kata Kunci : Analisis Sentimen, YouTube, Komentar, BERT, Ruangguru, Natural Language Processing, *Clash of Champions*.

ABSTRACT

The sentiment analysis technique aims to understand public opinions through textual data, including YouTube comments. These comments often reflect diverse perspectives from the audience regarding specific content. This research examines sentiment analysis of YouTube comments on the Clash of Champions video from Ruangguru's channel by utilizing the Bidirectional Encoder Representations from Transformers (BERT) model. The primary goal of this study is to identify and analyze sentiments within these comments using the BERT model, renowned for its advanced capability in comprehending natural language. The dataset used comprises YouTube comments that have been pre-processed, manually labeled, and supplemented with data augmentation techniques to ensure the quality and reliability of the analysis. The implementation employs the indobenchmark/indobert-base-p1 library of the BERT model and evaluates its performance using metrics such as accuracy, precision, recall, and F1-score. The analysis demonstrates that the model successfully achieves an accuracy of up to 96% in categorizing sentiments into four groups: positive, negative, neutral, and others. These findings provide Ruangguru with valuable insights into public sentiment and offer strategic recommendations for enhancing the quality of its content.

Keywords: *Sentiment Analysis, YouTube, Comments, BERT, Ruangguru, Natural Language Processing, Clash of Champions.*

LEMBAR PERSEMBAHAN

Puji dan syukur penulis panjatkan kepada Allah SWT atas limpahan rahmat dan karunia-Nya yang telah memberikan kemampuan kepada penulis untuk menyelesaikan Tugas Akhir ini. Penulis juga ingin mengucapkan terima kasih yang sebesar-besarnya kepada :

1. Allah SWT sebagai pencipta seluruh alam, telah memberikan kesehatan dan kelancaran kepada penulis dalam mengerjakan penelitian ini.
2. Kedua orang tua penulis Almarhum Bapak Hery Mardiono dan Ibu Linda Resmiati yang selalu menjadi sumber kekuatan, doa, dan kasih sayang yang tiada henti. Dan memberikan inspirasi , motivasi serta dukungan sepenuhnya kepada penulis dalam mengerjakan penelitian ini
3. Dosen pembimbing penulis Bapak Christian Sri Kusuma Aditya S.Kom., M.Kom dan Bapak Didih Rizki Chandranegara S.kom., M.Kom yang telah meluangkan waktu untuk membimbing penulis sehingga dapat menyelesaikan penelitian tugas akhir ini.
4. Dosen wali penulis Bapak Luqman Hakim, S.Kom., M.Kom yang telah memberikan bimbingan sejak awal perkuliahan hingga saat ini
5. Dosen dan staf di Program Studi Informatika Universitas Muhammadiyah Malang telah memberikan pengetahuan selama masa kuliah dan memberikan bantuan yang sangat berarti kepada penulis.
6. Kepada diri saya sendiri yang selalu berusaha dan bekerja keras untuk tidak menyerah dan melakukan yang terbaik.
7. Seluruh teman-teman seperjuangan yang telah membantu saya dalam proses pengerjaan Tugas Akhir ini.

Malang, 15 Januari 2025

Muhammad Galieh Mardiono

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan penelitian tugas akhir yang berjudul :

“Analisis Sentimen Komentar YouTube pada Video Clash of Champions di Channel Ruangguru Menggunakan Algoritma BERT”

Penelitian ini disusun sebagai bagian dari pemenuhan tugas akhir guna memperoleh gelar Sarjana Informatika di Universitas Muhammadiyah Malang. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap komentar pada video edukasi di platform YouTube menggunakan model *Bidirectional Encoder Representations from Transformers* (BERT), dengan harapan dapat memberikan pemahaman yang lebih mendalam terhadap opini masyarakat.

Penulis menyadari bahwa karya ini masih memiliki kekurangan dan ruang untuk pengembangan lebih lanjut. Oleh karena itu, dengan kerendahan hati, penulis sangat mengharapkan masukan serta kritik yang membangun demi penyempurnaan di masa mendatang. Semoga penelitian ini dapat memberikan kontribusi positif bagi pembaca, institusi akademik, maupun masyarakat luas.

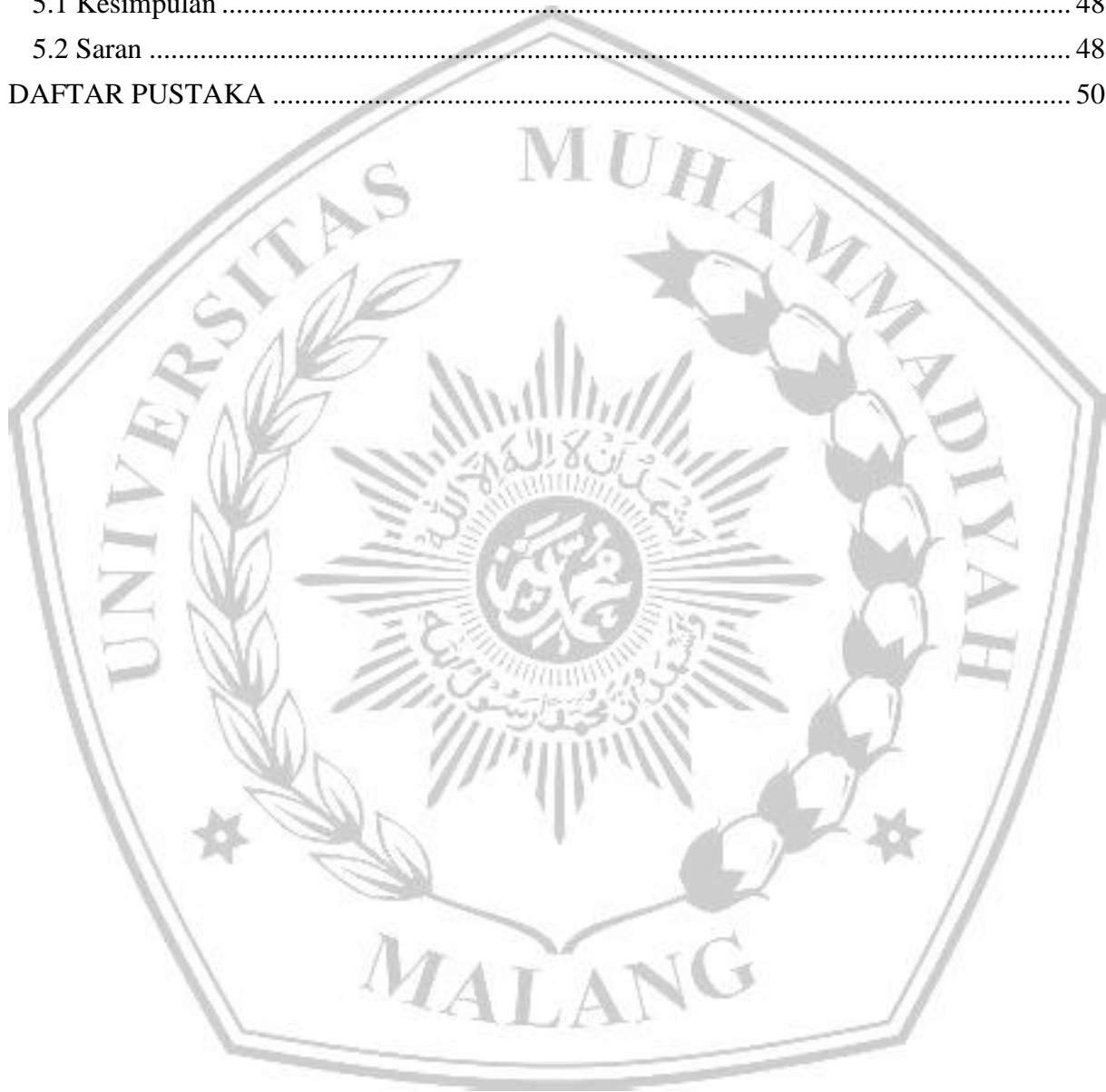
Malang, 15 Januari 2025

Muhammad Galieh Mardiono

DAFTAR ISI

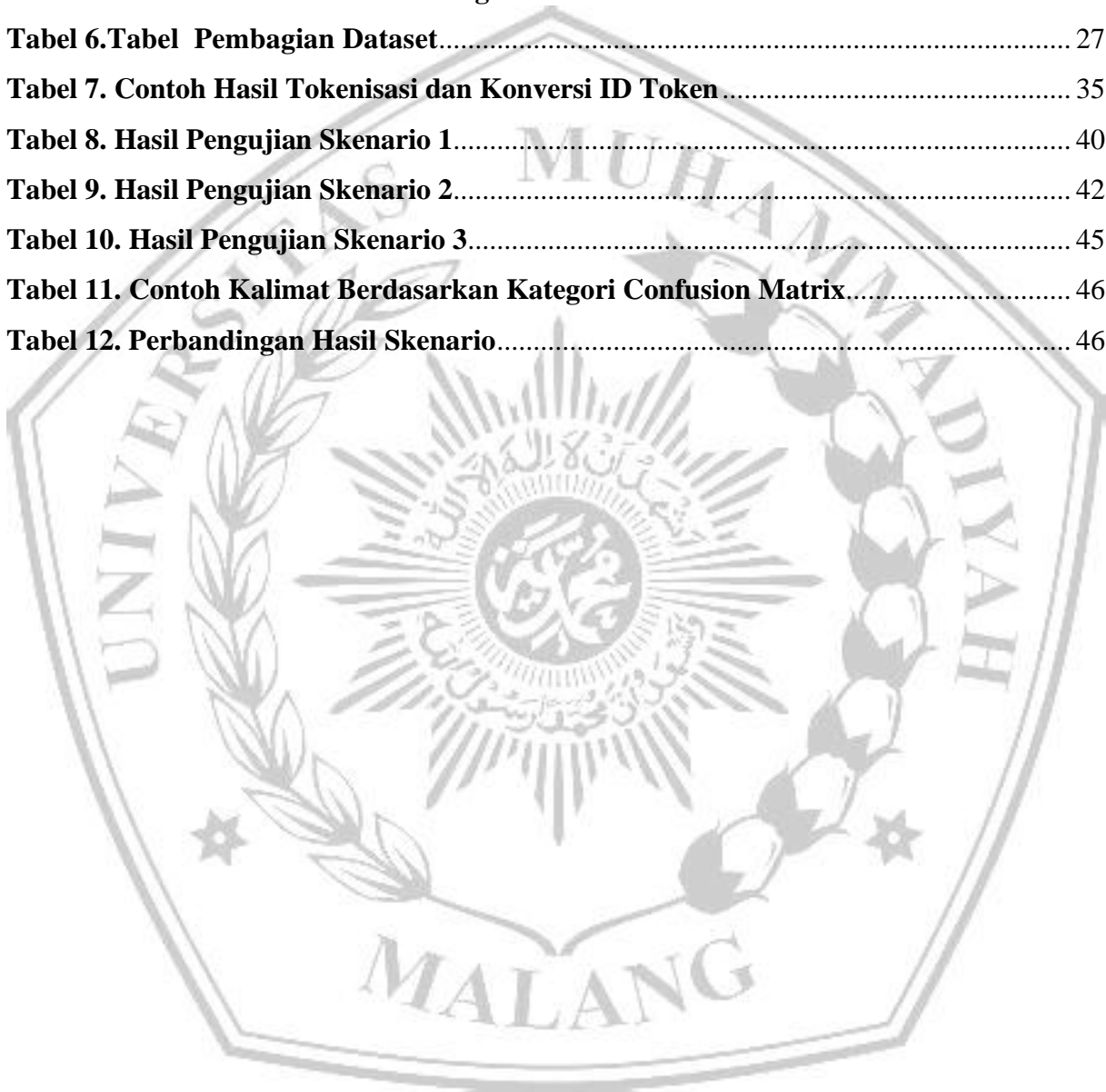
LEMBAR PERSETUJUAN.....	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERNYATAAN	iv
ABSTRAK	v
<i>ABSTRACT</i>	vi
LEMBAR PERSEMBAHAN	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN.....	13
1.1. Latar Belakang	13
1.2. Rumusan Masalah.....	15
1.3. Tujuan Penelitian.....	15
1.4. Batasan Masalah	15
BAB II TINJAUAN PUSTAKA.....	17
2.1 Studi Terkait.....	17
2.2 Analisis Sentimen	17
2.3 <i>Natural Language Processing</i> (NLP) dalam Analisis Sentimen	18
2.4 Algoritma BERT (<i>Bidirectional Encoder Representations from Transformers</i>)	18
2.4.1 Indobenchmark/indobert-base-p1	19
BAB III METODOLOGI PENELITIAN.....	22
3.1 Alur Penelitian	22
3.2 Pengumpulan Dataset.....	22
3.3 <i>Labelling</i> Data.....	23
3.3.1 Validasi Data Berlabel	24
3.4 Pre-Processing Dataset.....	25
3.5 <i>Data Split</i>	27
3.6 Augmentasi Dataset	28
3.7 Penerapan Model BERT	29
3.8 Evaluasi Model	31
BAB IV HASIL DAN PEMBAHASAN	34
4.1 Hasil Pengolahan dan Pre-Processing Data	34
4.2 Implikasi Hasil Analisis Sentimen.....	36

4.3 Hasil dan Evaluasi Pengujian Model	38
4.3.1 Skenario 1	38
4.3.2 Skenario 2	40
4.3.3 Skenario 3	43
4.4 Hasil Keseluruhan	46
BAB V KESIMPULAN	48
5.1 Kesimpulan	48
5.2 Saran	48
DAFTAR PUSTAKA	50



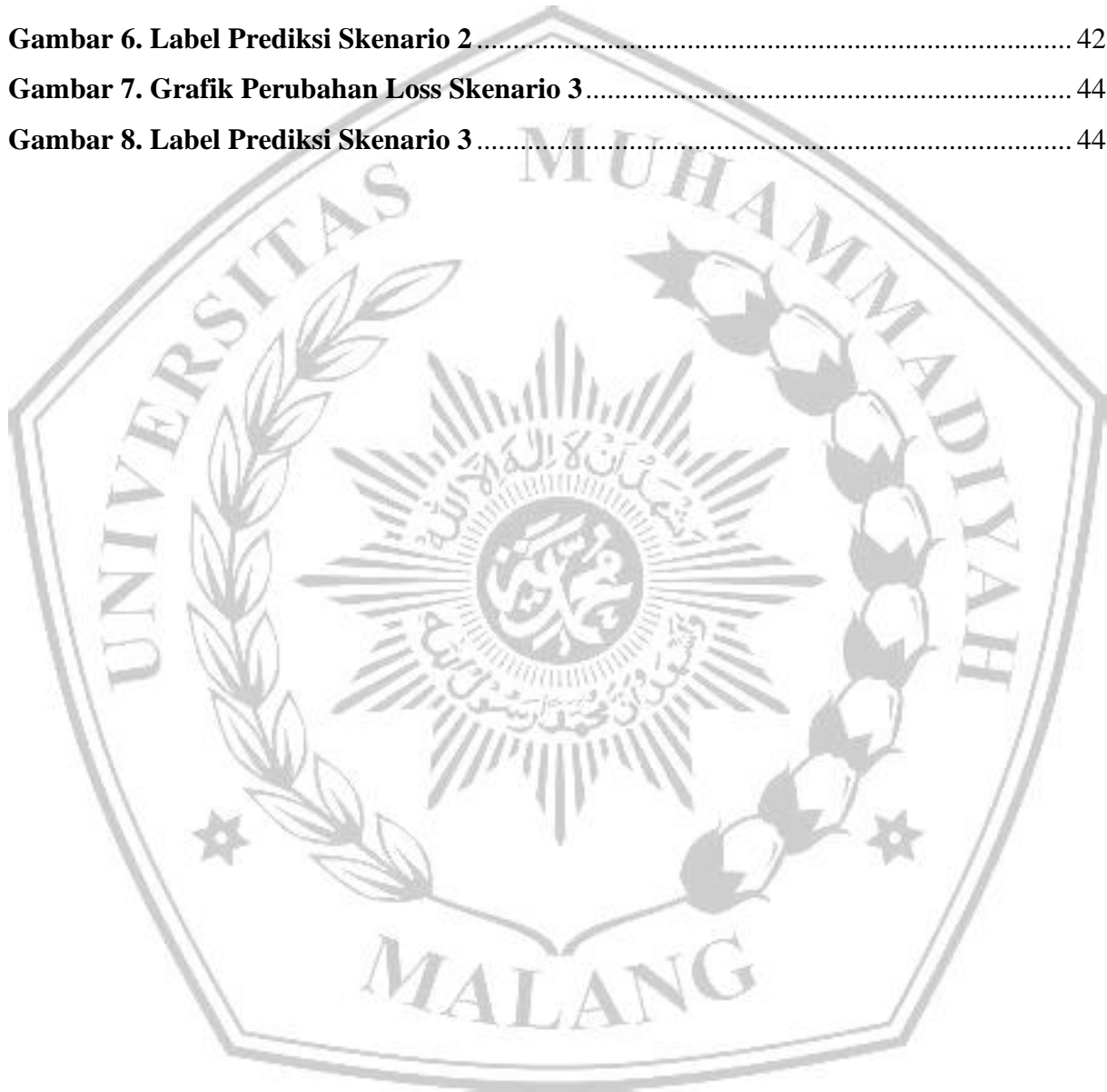
DAFTAR TABEL

Tabel 1. Tabel Contoh Hasil Pengumpulan Data.....	23
Tabel 2. Hasil Cohen Kappa Analysis	24
Tabel 3. Tabel Contoh Hasil Case Folding	25
Tabel 4. Tabel Contoh Hasil Data Cleaning	26
Tabel 5. Tabel Contoh hasil Stemming	27
Tabel 6. Tabel Pembagian Dataset.....	27
Tabel 7. Contoh Hasil Tokenisasi dan Konversi ID Token	35
Tabel 8. Hasil Pengujian Skenario 1.....	40
Tabel 9. Hasil Pengujian Skenario 2.....	42
Tabel 10. Hasil Pengujian Skenario 3.....	45
Tabel 11. Contoh Kalimat Berdasarkan Kategori Confusion Matrix.....	46
Tabel 12. Perbandingan Hasil Skenario.....	46



DAFTAR GAMBAR

Gambar 1. Alur Penelitian	22
Gambar 2. BERT Sequence Classification	29
Gambar 3. Grafik Perubahan Loss Skenario 1	39
Gambar 4. Label Prediksi Skenario 1	39
Gambar 5. Grafik Perubahan Loss Skenario 2	41
Gambar 6. Label Prediksi Skenario 2	42
Gambar 7. Grafik Perubahan Loss Skenario 3	44
Gambar 8. Label Prediksi Skenario 3	44



DAFTAR PUSTAKA

- [1] H. Al Rasyid Harpizon *et al.*, “Analisis Sentimen Komentar Di YouTube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naïve Bayes,” *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 1, 2022.
- [2] Laura Ceci, “Countries with the most YouTube users 2024,” Statista. Accessed: Aug. 10, 2024. [Online]. Available: <https://www.statista.com/statistics/280685/number-of-monthly-unique-youtube-users/>
- [3] R. Mas, R. W. Panca, K. Atmajal, and W. Yustanti², “Analisis Sentimen Customer Review Aplikasi Ruang Guru dengan Metode BERT (Bidirectional Encoder Representations from Transformers),” *JEISBI*, vol. 02, p. 2021.
- [4] D. Ayu Muthia, “ANALISIS SENTIMEN PADA REVIEW RESTORAN DENGAN TEKS BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE BAYES”, [Online]. Available: www.zomato.com
- [5] G. Saranya, G. Geetha, Chakrapani. K, M. K, and S. Karpagaselvi, “Sentiment analysis of healthcare Tweets using SVM Classifier,” in *2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, 2020, pp. 1–3. doi: 10.1109/ICPECTS49113.2020.9336981.
- [6] A. E. Putra and W. Maharani, “Depression Levels Detection Through Twitter Tweets Using RoBERTa Method,” *Journal of Information System Research (JOSH)*, vol. 3, no. 4, pp. 453–459, Jul. 2022, doi: 10.47065/josh.v3i4.1872.
- [7] S. Sumayah, F. Sembiring, and W. Jatmiko, “ANALYSIS OF SENTIMENT OF INDONESIAN COMMUNITY ON METAVERSE USING SUPPORT VECTOR MACHINE ALGORITHM,” vol. 4, no. 1, 2023, doi: 10.20884/1.jutif.2023.4.1.417.
- [8] H. M. Lee and Y. Sibaroni, “Comparison of IndoBERTweet and Support Vector Machine on Sentiment Analysis of Racing Circuit Construction in Indonesia,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 1, p. 99, Jan. 2023, doi: 10.30865/mib.v7i1.5380.
- [9] S. N. J. Fitriyyah, N. Safriadi, and E. E. Pratama, “Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 3, p. 279, Dec. 2019, doi: 10.26418/jp.v5i3.34368.
- [10] S. Helmiyah and A. Verdian, “JURNAL PENDIDIKAN STKIP ROSALIA LAMPUNG Analisis Sentimen Terhadap Minat Belajar pada Tayangan Acara CoC by Ruangguru Berdasarkan Tweets Menggunakan Metode NLP dan Model BERT.”
- [11] M. A. Devlin and B. P. Hayes, “Non-Intrusive Load Monitoring and Classification of Activities of Daily Living Using Residential Smart Meter Data,” *IEEE Transactions on Consumer Electronics*, vol. 65, no. 3, pp. 339–348, 2019, doi: 10.1109/TCE.2019.2918922.
- [12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>
- [13] A. Vaswani *et al.*, “Attention Is All You Need,” Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [14] C. Sun, X. Qiu, Y. Xu, and X. Huang, “How to Fine-Tune BERT for Text Classification?,” May 2019, [Online]. Available: <http://arxiv.org/abs/1905.05583>
- [15] N. A. R. Putri and Ardiansyah, “Analisis Sentimen Terhadap Kemajuan Kecerdasan Buatan di Indonesia Menggunakan BERT dan RoBERTa,” *Jurnal Sains dan Informatika*, vol. 9, no. 2, pp. 136–145, Nov. 2023, doi: 10.34128/jsi.v9i2.649.

- [16] Z. Shaheen, G. Wohlgenannt, and E. Filtz, "Large Scale Legal Text Classification Using Transformer Models," Oct. 2020.
- [17] A. Vaswani *et al.*, "Attention Is All You Need," Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [18] G. Hakim, T. N. Fatyanosa, and A. W. Widodo, "Analisis Sentimen Masyarakat terhadap Kereta Cepat Whoosh pada Platform X menggunakan IndoBERT," 2017. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [19] M. Widansyah, F. F. Az-zahra, and A. Pambudi, "Fine-Tuning Model Indobert (Indonesian Bidirectional Encoder Representations from Transformers) untuk Analisis Sentimen Berbasis Aspek pada Aplikasi M-Paspor," *Joutica*, 2024, [Online]. Available: <https://api.semanticscholar.org/CorpusID:276192199>
- [20] D. Sebagai *et al.*, "KLASIFIKASI SENTIMEN MENGGUNAKAN BIDIRECTIONAL LSTM DAN INDOBERT DENGAN DATASET TERBATAS TUGAS AKHIR."
- [21] D. Rudini, D. Gita Purnama, and A. Achmad Khan, "PENGUNAAN TEKNIK WEB SCRAPING DALAM APLIKASI PENGAMBILAN DATA DARI GOOGLE MAPS UNTUK MENUNJANG DIGITAL MARKETING," *Lentera: Multidisciplinary Studies*, vol. 2, no. 1, 2023, [Online]. Available: <https://lentera.publikasiku.id/index.php>
- [22] H. Al Jannah and D. Hermawan, "Analysis of Indonesian Society's Perceptions of the COVID-19 Vaccine in Youtube Comments Using Machine Learning Algorithms," in *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, 2022, pp. 72–77. doi: 10.1109/AiDAS56890.2022.9918796.
- [23] R. T. Handayanto and H. Herlawati, "Machine Learning Berbasis Desktop dan Web dengan Metode Jaringan Syaraf Tiruan Untuk Sistem Pendukung Keputusan," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 4, no. 1, pp. 15–26, Jun. 2020, doi: 10.31603/komtika.v4i1.3698.
- [24] A. Cahya Sutisna and N. Aini Rakhmawati, "ANALISIS SENTIMEN PENGGUNAAN PLATFORM CLOUD STORAGE PADA MAHASISWA DENGAN METODE NAIVE BAYES (STUDI KASUS PADA MAHASISWA SARJANA SISTEM INFORMASI ITS)," 2024.
- [25] N. Dwi Putranti and E. Winarko, "Analisis Sentimen Twitter untuk Teks Berbahasa Indonesia dengan Maximum Entropy dan Support Vector Machine," *IJCCS*, vol. 8, no. 1, pp. 91–100, 2014.
- [26] D. Alita and A. Rahman, "Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier," 2020.
- [27] A. Sri Widagdo *et al.*, "Analisis sentimen terhadap pelayanan Kesehatan berdasarkan ulasan Google Maps menggunakan BERT".
- [28] T. E. Putri and G. Ramadhan, "Penerapan Chatbot sebagai Alat Pembelajaran untuk Pengembangan Pendidikan Karakter," 2024.
- [29] J. U. S. Lazuardi and A. Juarna, "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI JOOX PADA ANDROID MENGGUNAKAN METODE BIDIRECTIONAL ENCODER REPRESENTATION FROM TRANSFORMER (BERT)," *Jurnal Ilmiah Informatika Komputer*, vol. 28, no. 3, pp. 251–260, 2023, doi: 10.35760/ik.2023.v28i3.10090.
- [30] I. Athiyyah Rahma and L. Hulliyyatus Suadaa, "PENERAPAN TEXT AUGMENTATION UNTUK MENGATASI DATA YANG TIDAK SEIMBANG PADA KLASIFIKASI TEKS BERBAHASA INDONESIA STUDI KASUS: DETEKSI JUDUL CLICKBAIT DAN KOMENTAR HATE SPEECH PADA BERITA ONLINE", doi: 10.25126/jtiik.2023107325.

- [31] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.04805>
- [32] T. Xiao and J. Zhu, "Introduction to Transformers: an NLP Perspective," Nov. 2023, [Online]. Available: <http://arxiv.org/abs/2311.17633>
- [33] M. Widiandyah, F. Frazna Az-Zahra, and A. Pambudi, "Fine-Tuning Model Indobert (Indonesian Bidirectional Encoder Representations from Transformers) untuk Analisis Sentimen Berbasis Aspek pada Aplikasi M-Paspor." [Online]. Available: <https://jurnalteknik.unisla.ac.id/index.php/informatika>





FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Muhammad Galieh Mardiono
 NIM : 202010370311442
 Judul TA : Analisis Sentimen Komentar YouTube Pada Video Clash of
 Champions di Channel Ruangguru Menggunakan Algoritma
 BERT

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	2%
2.	Bab 2 – Daftar Pustaka	25 %	8%
3.	Bab 3 – Analisis dan Perancangan	25 %	7%
4.	Bab 4 – Implementasi dan Pengujian	15 %	7%
5.	Bab 5 – Kesimpulan dan Saran	5 %	0%
6.	Makalah Tugas Akhir	20%	2%

*) Hasil cek plagiarism diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staff TU)


 (.....deny.....)