

**ANALISIS SENTIMEN JUDUL BERITA PADA MEDIA
MASSA MENGGUNAKAN ROBERTA**

TUGAS AKHIR



**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2026**

**ANALISIS SENTIMEN JUDUL BERITA PADA MEDIA
MASSA MENGGUNAKAN ROBERTA**

Tugas Akhir

Skema Skripsi

Tugas Akhir diajukan sebagai salah satu persyaratan memperoleh gelar Sarjana

Teknik pada Program Studi Teknik Informatika



Muhamad Rivai

NIM. 2101020112

**PROGRAM STUDI TEKNIK INFORMATIKA
JURUSAN TEKNIK ELEKTRO DAN INFORMATIKA
FAKULTAS TEKNIK DAN TEKNOLOGI KEMARITIMAN
UNIVERSITAS MARITIM RAJA ALI HAJI
TANJUNGPINANG
2026**

PERNYATAAN MENGENAI TUGAS AKHIR DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tugas akhir berjudul “Analisis Sentimen Judul Berita pada Media Massa Menggunakan RoBERTa” adalah benar karya saya dengan arahan dari komisi pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Apabila di kemudian hari terbukti pernyataan saya tidak benar dan melanggar peraturan yang berlaku dalam karya tulis dan hak intelektual, maka saya bersedia ijazah yang telah saya terima untuk ditarik kembali oleh Universitas Maritim Raja Ali Haji dan menerima sanksi lainnya sesuai dengan peraturan yang berlaku. Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Maritim Raja Ali Haji.

Tanjungpinang, 10 Januari 2026



Muhamad Rivai
NIM. 2101020112



© Hak Cipta Milik Universitas Maritim Raja Ali Haji, Tahun 2026

Hak Cipta Dilindungi Undang- Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah; dan pengutipan tersebut tidak merugikan kepentingan Universitas Maritim Raja Ali Haji.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin Universitas Maritim Raja Ali Haji.



HALAMAN PERSETUJUAN

Judul : Analisis Sentimen Judul Berita pada Media Massa
Menggunakan RoBERTa
Nama : Muhamad Rivai
NIM : 2101020112
Program Studi : Teknik Informatika
Tanggal Persetujuan : 9 Desember 2025

Disetujui oleh,
Komisi Pembimbing


Tekad Matulatan, S.Sos., S.Kom.,
M.Info.Tech
(Ketua)

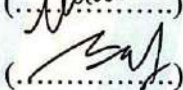

Nolan Efranda, M.Kom
(Anggota)

HALAMAN PENGESAHAN

Judul : Analisis Sentimen Judul Berita pada Media Massa Menggunakan RoBERTa
Nama : Muhamad Rivai
NIM : 2101020112

Telah berhasil dipertahankan di hadapan Komisi Penguji sebagai bagian persyaratan yang diperlukan dalam memperoleh gelar Sarjana Teknik pada Program Studi Teknik Informatika, Jurusan Teknik Elektro dan Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

KOMISI PENGUJI

Ketua	: Nurul Hayaty, S.T., M.Cs	()
Anggota I	: Novrizal Fattah Fahmitra, S.Kom., M.Kom	()
Anggota II	: Marisha Pertiwi, S.Tr.Kom., M.Kom	()
Anggota III	: Tekad Matulatan, S.Sos., S.Kom., M.Info.Tech	()
Anggota IV	: Nolan Efranda, M.Kom	()

Mengetahui,

Ketua Jurusan Teknik Elektro dan Informatika


Hollanda Arief Kusuma, S.IK., M.Si

NIP. 198904012019031016

Tanggal Ujian: 23 Desember 2025 Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul **“Analisis Sentimen Judul Berita pada Media Massa Menggunakan RoBERTa”**. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi Kemaritiman, Universitas Maritim Raja Ali Haji.

Selama proses penyusunan skripsi ini, penulis menyadari banyaknya kontribusi dari berbagai pihak yang telah memberikan dukungan dan arahan, baik secara langsung maupun tidak langsung. Untuk itu, pada kesempatan ini penulis menyampaikan rasa terima kasih yang mendalam kepada:

1. Prof. Dr. Agung Dhamar Syakti, S.Pi., DEA., selaku Rektor Universitas Maritim Raja Ali Haji.
2. Ibu Martaleli Bettiza, S.Si., M.Sc., selaku Dekan Fakultas Teknik dan Teknologi Kemaritiman Universitas Maritim Raja Ali Haji.
3. Bapak Hollanda Arief Kusuma, S.I.K., M.Si., selaku Ketua Jurusan Teknik Elektro dan Informatika.
4. Bapak Tekad Matulatan, S.Sos., S.Kom., M.Info.Tech., selaku Dosen Pembimbing I, yang dengan sabar telah memberikan arahan, bimbingan, serta ilmu yang bermanfaat dari awal hingga selesainya penulisan skripsi ini.
5. Bapak Nolan Efranda, M.Kom., selaku Dosen Pembimbing II, yang selalu meluangkan waktu untuk memberikan masukan dan motivasi selama penyusunan skripsi ini.
6. Seluruh dosen di Fakultas Teknik dan Teknologi Kemaritiman Universitas Maritim Raja Ali Haji yang telah membagikan ilmu, pengalaman, serta bimbingan selama masa perkuliahan.
7. Staf administrasi Fakultas Teknik dan Teknologi Kemaritiman Universitas Maritim Raja Ali Haji yang selalu siap membantu kelancaran proses akademik dan penelitian.
8. Kedua orang tua tercinta, yang selalu memberikan kasih sayang, doa, serta dukungan tanpa henti kepada penulis.

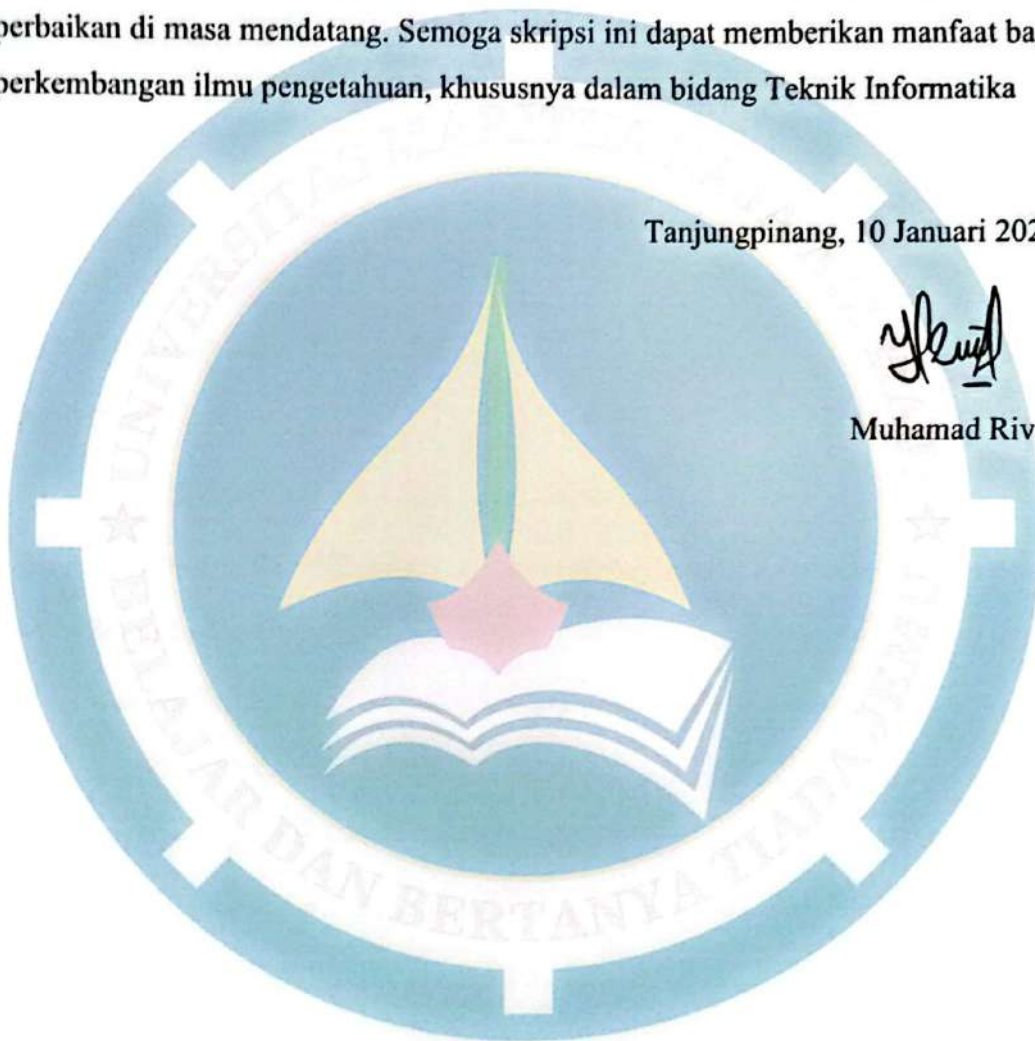
9. Seluruh anggota keluarga yang selalu memberikan motivasi dan dukungan sehingga penelitian ini terselesaikan.
10. Semua pihak yang tidak dapat penulis sebutkan satu per satu, namun telah memberikan dukungan dan bantuan hingga selesainya penulisan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna karena keterbatasan pengetahuan dan kemampuan yang dimiliki. Oleh karena itu, penulis dengan rendah hati mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi perkembangan ilmu pengetahuan, khususnya dalam bidang Teknik Informatika

Tanjungpinang, 10 Januari 2026



Muhamad Rivai



DAFTAR ISI

HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
ABSTRAK	vi
ABSTRACT.....	vii
PRAKATA.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Perumusan Masalah	4
C. Batasan Penelitian	4
D. Tujuan Penelitian	5
E. Manfaat Penelitian	5
BAB II KAJIAN LITERATUR	6
A. Tinjauan Pustaka	6
B. Landasan Teori.....	9
1. Analisis Sentimen	9
2. <i>Scraping</i>	10
3. Media Massa	11
4. <i>Data Preprocessing</i>	11
5. Pelabelan Data	13
6. <i>Crowdsourced Labeling</i>	14

7.	Statistik Deskriptif	14
8.	<i>K-Fold Cross Validation</i>	14
9.	Tokenizer	15
10.	RoBERTa.....	16
11.	<i>Confusion Matrix</i>	21
BAB III METODE PENELITIAN.....		24
A.	Waktu dan Tempat Penelitian	24
B.	Bahan dan Alat.....	24
C.	Prosedur Pelaksanaan Penelitian.....	25
D.	Perancangan Sistem	30
1.	Perancangan Pemodelan RoBERTa.....	30
2.	Perancangan Aplikasi	32
E.	Analisis Data	33
1.	★ <i>Data Preprocessing</i>	33
2.	Perhitungan Manual RoBERTa	34
3.	Pengujian	44
BAB IV HASIL DAN PEMBAHASAN		46
A.	Hasil	46
1.	Visualisasi Sentimen.....	46
2.	Hasil Evaluasi Model.....	48
3.	Hasil Implementasi	48
B.	Pembahasan.....	49
1.	<i>Preprocessing</i> dan Pelabelan Dataset	49
2.	Pembagian Dataset.....	60
3.	Implementasi RoBERTa	61
4.	Hasil Pelatihan Model.....	62

5. Evaluasi Model pada Data <i>Testing</i>	68
BAB V SIMPULAN DAN SARAN	71
A. Simpulan	71
B. Saran.....	71
DAFTAR PUSTAKA	72
LAMPIRAN.....	82



DAFTAR TABEL

Tabel 1. Keaslian Penelitian.....	8
Tabel 2. <i>Confusion Matrix</i> Dua kelas	22
Tabel 3. <i>Confusion Matrix</i> Tiga Kelas.....	22
Tabel 4. Tahap <i>Cleaning</i> pada Perhitungan Manual.....	34
Tabel 5. Tahap <i>Case Folding</i> pada Perhitungan Manual.....	34
Tabel 6. Hasil Konversi Token ID	35
Tabel 7. Hasil <i>Token Embedding</i>	35
Tabel 8. Hasil <i>Position Embedding</i>	36
Tabel 9. Hasil Penjumlahan <i>Token Embedding</i> dan <i>Position Embedding</i>	36
Tabel 10. Hasil Perhitungan Matriks QKV	38
Tabel 11. Hasil Perhitungan <i>Attention Scores</i>	39
Tabel 12. Hasil Eksponen Tiap <i>Attention Scores</i> Baris Pertama	39
Tabel 13. Hasil <i>Softmax Attention Scores</i> untuk Baris Pertama	40
Tabel 14. Hasil <i>Softmax</i> untuk Semua <i>Attention Scores</i>	40
Tabel 15. Hasil Perhitungan Matriks V untuk Token Lain.....	41
Tabel 16. Hasil <i>Attention Output</i> untuk Semua Data.....	41
Tabel 17. Hasil <i>Softmax</i> dari Logits.....	43
Tabel 18. <i>Data Dummy</i> untuk Pengujian	44
Tabel 19. Hasil <i>Dummy</i> Pengujian dengan Model RoBERTa.....	45
Tabel 20. <i>Confusion Matrix</i> Pengujian Model Roberta dengan <i>Data Dummy</i>	45
Tabel 21. Tahap <i>Cleaning</i> Pra-Penelitian	49
Tabel 22. Tahap <i>Case Folding</i> Pra-Penelitian	50
Tabel 23. Tabel Normalisasi Kata Pra-Penelitian	50
Tabel 24. Tahap Tokenisasi Pra-Penelitian.....	51
Tabel 25. Tahap <i>Stopword Removal</i> Pra-penelitian.....	51
Tabel 26. Tahap <i>Stemming</i> Pra-Penelitian	52
Tabel 27. Data Final Tahap Pra-Penelitian	52
Tabel 28. Hasil Pelabelan IndoBERT Pra-Penelitian	53
Tabel 29. Hasil Pelabelan <i>Crowdsourced</i> Pra-Penelitian	55
Tabel 30. Hasil Percobaan Pra-Penelitian.....	56
Tabel 31. Tahap <i>Cleaning</i> pada Dataset	57

Tabel 32. Tahap <i>Case Folding</i> pada Dataset	58
Tabel 33. Hasil pelabelan IndoBERT pada Dataset.....	58
Tabel 34. Pembagian Dataset.....	60
Tabel 35. Metrik Evaluasi Model	69
Tabel 36. Metrik Evaluasi Rata-Rata Model	69



DAFTAR GAMBAR

Gambar 1. Arsitektur Model RoBERTa	17
Gambar 2. Prosedur Pelaksanaan Penelitian	26
Gambar 3. Grafik Frekuensi Data Berdasarkan Sumbernya	28
Gambar 4. Grafik Distribusi Jumlah Kata Per Judul	28
Gambar 5. Perancangan Pemodelan RoBERTa	31
Gambar 6. <i>Flowchart</i> Aplikasi	32
Gambar 7. <i>User Interface Input Manual</i>	32
Gambar 8. <i>User Interface Input Batch</i>	33
Gambar 9. <i>Word Cloud</i> Sentimen Negatif	46
Gambar 10. <i>Word Cloud</i> Sentimen Netral	47
Gambar 11. <i>Word Cloud</i> Sentimen Positif	47
Gambar 12. Hasil Implementasi <i>User Interface Input Manual</i>	48
Gambar 13. Hasil Implementasi <i>User Interface Input Batch</i>	49
Gambar 14. Dokumentasi Model IndoBERT	53
Gambar 15. Grafik Frekuensi Sentimen Hasil IndoBERT	54
Gambar 16. <i>Screenshot</i> Hasil Pelabelan <i>Crowdsourced</i>	55
Gambar 17. Grafik Frekuensi Sentimen Hasil Pelabelan <i>Crowdsourced</i>	56
Gambar 18. Grafik Hasil Pelabelan IndoBERT pada Dataset	59
Gambar 19. Grafik Persebaran Sentimen pada Dataset	60
Gambar 20. Grafik Perbandingan Utilitas dan <i>Learning Rate</i>	63
Gambar 21. Grafik Perbandingan Waktu <i>Training</i> dan <i>Learning Rate</i>	64
Gambar 22. Grafik Perbandingan Akurasi dan <i>Learning Rate</i>	65
Gambar 23. Grafik <i>Training</i> dan <i>Validation Loss Model</i>	66
Gambar 24. Grafik <i>Training</i> dan <i>Validation Accuracy Model</i>	67
Gambar 25. <i>Confusion Matrix</i> Model RoBERTa	68
Gambar 26. Perbandingan Frekuensi Sentimen Asli dan Sentimen Prediksi	70
Gambar 27. Hasil Pelabelan IndoBERT <i>Full Preprocessing</i>	83
Gambar 28. Hasil Pelabelan IndoBERT <i>Non-Full Preprocessing</i>	83
Gambar 29. Hasil Pelabelan <i>Crowdsourced Full Preprocessing</i>	83
Gambar 30. Hasil Pelabelan <i>Crowdsourced Non-Full Preprocessing</i>	83

DAFTAR LAMPIRAN

Lampiran 1. Landasan Teori Berbentuk Persamaan	82
Lampiran 2. <i>Confusion Matrix</i> Hasil Percobaan Model Pra-Penelitian.....	83
Lampiran 3. Sampel Data <i>Testing</i> Hasil Pengujian Model RoBERTa	84

