

**PERBANDINGAN INDOBERT DAN *HYBRID* INDOBERT-CNN
UNTUK ANALISIS SENTIMEN KOMENTAR YOUTUBE
TENTANG PENERAPAN KEMBALI UJIAN NASIONAL**

TUGAS AKHIR



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

**PERBANDINGAN INDOBERT DAN *HYBRID* INDOBERT-CNN
UNTUK ANALISIS SENTIMEN KOMENTAR YOUTUBE
TENTANG PENERAPAN KEMBALI UJIAN NASIONAL**

Tugas Akhir

Skema Skripsi

Tugas Akhir diajukan sebagai salah satu persyaratan memperoleh gelar Sarjana

Teknik pada Program Studi Teknik Informatika



KELVIN AMZANI PASHA

NIM. 2101020091

**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 “Perbandingan IndoBERT Dan *Hybrid* IndoBERT-CNN Untuk Analisis Sentimen Komentar Youtube Tentang Penerapan Kembali Ujian Nasional” 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 Juni 2026



Kelvin amzani pasha
NIM 2101020091



© 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 : Perbandingan IndoBERT Dan *Hybrid* IndoBERT-
CNN Untuk Analisis Sentimen Komentar Youtube
Tentang Penerapan Kembali Ujian Nasional
Nama : Kelvin Amzani Pasha
NIM : 2101020091
Program Studi : Teknik Informatika
Tanggal Persetujuan : 10 Juni 2026

Disetujui oleh,
Komisi Pembimbing



Nerfita Nikentari, S.T., M.Cs.
(Ketua)



Novrizal Fatah Fahmitra, S.kom.,
M.Kom.
(Anggota)

HALAMAN PENGESAHAN

Judul : Perbandingan IndoBERT Dan Hybrid
IndoBERT-CNN Untuk Analisis Sentimen
Komentar Youtube Tentang Penerapan Kembali
Ujian Nasional

Nama : Kelvin Amzani Pasha

NIM : 2101020091

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

KOMISI PENGUJI

Ketua	: Nola Ritha, S.T., M.Cs	(.....)
Anggota I	: Feri Irawan, S.kom., M.Kom	(.....)
Anggota II	: Rifaldi Herikson, M.Kom	(.....)
Anggota III	: Nerfita Nikentari, S.T., M.Cs.	(.....)
Anggota IV	: Novrizal Fattah Fahmitra, S.kom., M.Kom.	(.....)

Mengetahui,

Ketua Jurusan Teknik Elektro dan Informatika


Hollanda Arief Kusuma, S.IK., M.Si
NIP. 198904012019031016

Tanggal Ujian: 24 Juni 2026

Tanggal Lulus: 24 Juni 2026

PRAKATA

Puji dan syukur dipanjatkan kepada Allah *subhanahu wa ta'ala* atas segala karunia-Nya sehingga laporan tugas akhir ini berhasil diselesaikan. Tema yang dipilih dalam tugas akhir yang akan dilaksanakan pada bulan Juni 2026 ini ialah konsumsi daya, dengan judul **“Perbandingan IndoBERT Dan Hybrid IndoBERT-CNN Untuk Analisis Sentimen Komentar Youtube Tentang Penerapan Kembali Ujian Nasional”**.

Terima kasih penulis ucapkan kepada Ibu Nerfita Nikentari, S.T., M.Cs. dan Bapak Novrizal Fattah Fahmitra, S.kom., M.Kom. yang telah banyak memberi saran. Ungkapan terima kasih juga disampaikan kepada ayah, ibu, serta seluruh keluarga, atas segala doa dan kasih sayang.

Semoga laporan tugas akhir ini bermanfaat.

Tanjungpinang, 10 Juni 2026



Kelvin Amzani Pasha

DAFTAR ISI

HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN.....	v
ABSTRAK	vi
<i>ABSTRACT</i>	vii
PRAKATA.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Perumusan Masalah	2
C. Batasan Penelitian	2
D. Tujuan Penelitian	3
E. Manfaat Penelitian	3
BAB II KAJIAN LITERATUR	5
A. Tinjauan Pustaka	5
B. Landasan Teori.....	12
1. Analisis Sentimen	12
2. NLP (<i>Natural Language Processing</i>)	12
3. <i>Preprocessing</i> Teks.....	13
4. <i>Lexicon-Based</i>	14
5. <i>Transformer</i>	15
6. IndoBERT	17
7. CNN (<i>Convolutional Neural Network</i>).....	19

8.	<i>Model Hybrid</i>	21
9.	<i>Fully Connected Layer (FC Layer)</i>	21
10.	<i>Softmax</i>	22
11.	<i>Confusion Matrix</i>	22
12.	<i>Hold-Out Validation</i>	23
BAB III METODE PENELITIAN		24
A.	Waktu dan Tempat Penelitian	24
B.	Bahan dan Alat	24
C.	Pengumpulan Data	24
D.	Prosedur Pelaksanaan Penelitian	25
E.	Perancangan Sistem	29
1.	Perancangan Model IndoBERT	29
2.	Perancangan Model <i>Hybrid</i> (IndoBERT + CNN)	30
F.	Analisis Data	31
1.	Perhitungan Manual IndoBERT	31
2.	Perhitungan Manual <i>Hybrid</i> (IndoBERT + CNN)	34
BAB IV HASIL DAN PEMBAHASAN		38
A.	Data Penelitian	38
B.	Hasil <i>Preprocessing</i>	39
1.	<i>Case Folding</i>	39
2.	<i>Filtering</i>	39
3.	<i>Normalization</i>	40
4.	<i>Tokenization</i>	40
5.	<i>Stopword Removal</i>	41
6.	<i>Stemming</i>	42
C.	Hasil Pelabelan	42

D. Hasil Pelatihan Model IndoBERT	45
E. Hasil Pelatihan Model <i>Hybrid</i>	47
F. Perbandingan Best Model	48
G. Analisis Sentimen Publik dan Pembahasan Hasil.....	49
1. <i>Classification Report</i> data validation.....	50
2. <i>Confusion Matrix</i> data validation.....	50
3. <i>Classification Report</i> data testing	51
4. <i>Confusion Matrix</i> data testing	52
5. Validasi Prediksi Model terhadap Penilaian Ahli Bahasa.....	53
6. Interpretasi Hasil	57
H. Implementasi Aplikasi Streamlit.....	59
1. Dashboard	59
2. Evaluasi	60
3. Tentang.....	60
BAB V SIMPULAN DAN SARAN.....	61
A. Simpulan	61
B. Saran.....	62
DAFTAR PUSTAKA	63
LAMPIRAN.....	67

DAFTAR TABEL

Tabel 1. Penelitian yang pernah dilakukan	7
Tabel 2. Teknologi Dasar NLP	12
Tabel 3. Contoh Preprocessing Teks (Sumber : [20]).....	14
Tabel 4. Confusion Matrix Dua Kelas	23
Tabel 5. Alat penelitian	24
Tabel 6. Hasil tokenisasi model IndoBERT	32
Tabel 7. Hasil <i>embedding</i> model IndoBERT	32
Tabel 8. Data untuk uji perhitungan bobot dan bias model IndoBERT	32
Tabel 9. Bobot dan bias model IndoBERT	32
Tabel 10. Perhitungan menggunakan bobot dan bias model IndoBERT	33
Tabel 11. Hasil perhitungan <i>softmax</i> model IndoBERT	34
Tabel 12. Hasil tokenisasi model <i>hybrid</i>	34
Tabel 13. Hasil <i>embedding</i> model <i>hybrid</i>	35
Tabel 14. Data untuk uji perhitungan bobot dan bias model <i>hybrid</i>	35
Tabel 15. Tabel bobot dan bias model <i>hybrid</i>	36
Tabel 16. Perhitungan menggunakan bobot dan bias model <i>hybrid</i>	36
Tabel 17. Hasil perhitungan <i>softmax</i> model <i>hybrid</i>	37
Tabel 18. Dataset hasil <i>cleaning</i>	38
Tabel 19. Dataset hasil <i>case folding</i>	39
Tabel 20. Dataset hasil <i>filtering</i>	40
Tabel 21. Dataset hasil <i>normalization</i>	40
Tabel 22. Dataset hasil <i>tokenization</i>	41
Tabel 23. Dataset hasil <i>stopword removal</i>	41
Tabel 24. Dataset hasil <i>stemming</i>	42
Tabel 25. Dataset hasil pelabelan.....	43
Tabel 26. Pelatihan Model IndoBERT dengan <i>grid search</i>	46
Tabel 27. Log pelatihan model IndoBERT	46
Tabel 28. Pelatihan model <i>hybrid</i> dengan <i>grid search</i>	47
Tabel 29. Log pelatihan model <i>hybrid</i>	48
Tabel 30. Perbandingan evaluasi kedua model	49
Tabel 31. <i>Classification Report</i> data validation model IndoBERT	50

Tabel 32. <i>Classification Report</i> data validation model <i>hybrid</i>	50
Tabel 33. <i>Classification Report</i> data testing model IndoBERT.....	51
Tabel 34. <i>Classification Report</i> data testing model <i>hybrid</i>	51
Tabel 35. Perbandingan hasil pelabelan manual dan aplikasi sentimen	54



DAFTAR GAMBAR

Gambar 1. Arsitektur Transformer (Sumber : [22]).....	16
Gambar 2. Model Attention (Sumber : [22])	17
Gambar 3. pre-training and fine-tuning BERT (Sumber : [25])	18
Gambar 4. BERT Embedding (Sumber : [25])	19
Gambar 5. Arsitektur lapisan CNN.....	20
Gambar 6. Kerangka Penelitian	26
Gambar 7. <i>Flowchart Preprocessing Data</i>	27
Gambar 8. Diagram alir metode IndoBERT	29
Gambar 9. Diagram alir metode <i>hybrid</i> (IndoBERT + CNN)	30
Gambar 10. Grafik distribusi sentimen	44
Gambar 11. Grafik pelabelan data <i>training</i>	44
Gambar 12. Grafik pelabelan data <i>validation</i>	45
Gambar 13. Grafik pelabelan data <i>testing</i>	45
Gambar 14. Grafik evaluasi validasi model IndoBERT	47
Gambar 15. Grafik evaluasi validasi model <i>Hybrid</i>	48
Gambar 16. Grafik perbandingan evaluasi kedua model.....	49
Gambar 17. <i>Confusion Matrix data validation</i> IndoBERT	50
Gambar 18. <i>Confusion Matrix data validation Hybrid</i>	51
Gambar 19. <i>Confusion Matrix data testing</i> IndoBERT.....	52
Gambar 20. <i>Confusion Matrix data testing hybrid</i>	52
Gambar 21. Tampilan halaman menu dashboard.....	59
Gambar 22. Tampilan halaman menu evaluasi	60
Gambar 23. Tampilan halaman menu tentang	60

DAFTAR LAMPIRAN

Lampiran 1. Data Mentah Hasil <i>Scrapping</i> Youtube.....	67
Lampiran 2. Dataset Hasil Pelabelan dan Pembagian Data.....	68
Lampiran 3. Hasil <i>Grid Search</i> Kedua Model.....	69
Lampiran 4. Form Penilaian Aplikasi Skripsi Oleh Ahli Bahasa.....	70

