



SKRIPSI

KLASIFIKASI EMOSI PADA TEKS BAHASA INDONESIA DAN EMOJI MENGGUNAKAN ALGORITMA HYBRID CNN-BILSTM

RANGGA WIDIASMARA

NPM 21081010085

DOSEN PEMBIMBING

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU

Chrystia Aji Putra, S.Kom., MT.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**KLASIFIKASI EMOSI PADA TEKS BAHASA
INDONESIA DAN EMOJI MENGGUNAKAN
ALGORITMA HYBRID CNN-BILSTM**

RANGGA WIDIASMARA

NPM 21081010085

DOSEN PEMBIMBING

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU

Chrystia Aji Putra, S.Kom., MT.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

Halaman ini sengaja dikosongkan

LEMBAR PENGESAHAN

KLASIFIKASI EMOSI PADA TEKS BAHASA INDONESIA DAN EMOJI MENGUNAKAN ALGORITMA HYBRID CNN-BILSTM

Oleh :
RANGGA WIDIASMARA
NPM. 21081010085

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 15 Juli 2025.

Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
NIP. 19700619 202121 1 009


.....
(Pembimbing I)

Chrystia Aji Putra, S.Kom., MT.
NIP. 19861008 2021211 001


.....
(Pembimbing II)

Made Hanindia Prami Swari, S.Kom, M.Cs.
NIP. 19890205 201803 2 001


.....
(Ketua Penguji)

Fawwaz Ali Akbar, S.Kom, M.Kom.
NIP. 19920317 201803 1 002


.....
(Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

KLASIFIKASI EMOSI PADA TEKS BAHASA INDONESIA DAN EMOJI MENGUNAKAN ALGORITMA HYBRID CNN-BILSTM

Oleh :
RANGGA WIDIASMARA
NPM. 21081010085



Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer

Fetty Tri Anggraeny, S.Kom. M.Kom
NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini:

Nama : Rangga Widiasmara
NPM : 21081010085
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Tugas Akhir/Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi/Tesis/Desertasi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 22 Juli 2025
Yang Membuat pernyataan


Rangga Widiasmara
NPM. 21081010085

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa /NPM : Rangga Widiasmara / 21081010085
Judul Skripsi : Klasifikasi Emosi pada Teks Bahasa Indonesia dan
Emoji menggunakan Algoritma Hybrid CNN-
BiLSTM
Dosen Pembimbing : 1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
2. Chrystia Aji Putra, S.Kom., MT.

Penelitian ini merancang dan mengevaluasi secara komparatif empat arsitektur model *deep learning* untuk mengatasi tantangan klasifikasi emosi pada teks dan emoji berbahasa Indonesia yang kompleks dan ambigu. Dengan membandingkan pendekatan *hybrid* (serial vs. paralel) dari *Convolutional Neural Network* (CNN) dan *Bidirectional Long Short-Term Memory* (BiLSTM) serta skema input (tunggal vs. ganda) pada dataset *Emotion Twitter* dari IndoNLU, penelitian ini menemukan bahwa pemrosesan fitur emoji secara eksplisit melalui skema input ganda secara signifikan meningkatkan kinerja. Model terbaik, yaitu arsitektur *hybrid* paralel dengan input ganda, berhasil mencapai akurasi pengujian sebesar 90,16% dan *Macro F1-Score* 0,90. Model ini terbukti efektif dalam menangani kelas ambigu seperti 'takut' dan 'sedih', meskipun tantangan yang persisten tetap ada dalam membedakan kelas 'senang' dan 'sedih' akibat kompleksitas linguistik. Hasil ini menyimpulkan bahwa pendekatan *hybrid* CNN-BiLSTM dengan arsitektur paralel dan multi-input merupakan metode yang sangat efektif dan menjanjikan untuk analisis emosi pada konten digital berbahasa Indonesia.

Kata kunci: Klasifikasi Emosi, Hybrid CNN-BiLSTM, Pemrosesan Bahasa Alami, Analisis Sentimen, Teks dan Emoji, Deep Learning

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Rangga Widiasmara / 21081010085
Thesis Title : Emotion Classification in Indonesian Text and
Emoji using Hybrid CNN-BiLSTM Algorithm
Advisor : 1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU
2. Chrystia Aji Putra, S.Kom., MT.

This research designs and comparatively evaluates four deep learning model architectures to address the challenges of emotion classification in complex and ambiguous Indonesian text and emoji. By comparing hybrid approaches (serial vs. parallel) of Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) as well as input schemes (single vs. dual) on the Emotion Twitter dataset from IndoNLU, this study finds that explicitly processing emoji features through a dual-input scheme significantly improves performance. The best model, a parallel hybrid architecture with dual inputs, successfully achieved a test accuracy of 90.16% and a Macro F1-Score of 0.90. This model proved effective in handling ambiguous classes like 'fear' and 'sadness', although a persistent challenge remains in distinguishing between 'happy' and 'sadness' classes due to linguistic complexities. These results conclude that the hybrid CNN-BiLSTM approach with a parallel and multi-input architecture is a highly effective and promising method for emotion analysis in Indonesian digital content.

Keywords: Emotion Classification, Hybrid CNN-BiLSTM, Natural Language Processing, Sentiment Analysis, Text and Emoji, Deep Learning

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Klasifikasi Emosi Pada Teks Bahasa Indonesia Dan Emoji Menggunakan Algoritma Hybrid CNN-BiLSTM” dengan baik. Penyusunan skripsi ini diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penyelesaian skripsi ini tidak terlepas dari bimbingan, dukungan, dan bantuan dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis ingin menyampaikan rasa terima kasih yang tulus dan penghargaan setinggi-tingginya kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.MT., selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom., M.Kom., selaku Koordinator Program Studi Informatika.
4. Bapak Dr. Ir. I Gede Susrama Mas Diyasa, S.T., M.T., IPU, selaku Dosen Pembimbing I, yang dengan penuh kesabaran dan dedikasi telah memberikan wawasan mendalam, motivasi, serta arahan yang sangat berarti dalam setiap tahap penyusunan skripsi ini.
5. Bapak Chrystia Aji Putra, S.Kom., M.T., selaku Dosen Pembimbing II, yang telah meluangkan waktu untuk diskusi yang konstruktif, serta memberikan bimbingan teknis dan masukan berharga yang sangat membantu dalam pengembangan sistem dan penyempurnaan laporan.

6. Ibu Made Hanindia Prami Swari, S.Kom., M.Cs., selaku Dosen Penguji I, atas evaluasi kritis, saran yang tajam, dan masukan yang sangat membangun untuk penyempurnaan skripsi ini.
7. Bapak Fawwaz Ali Akbar, S.Kom., M.Kom., selaku Dosen Penguji II, atas arahan, pertanyaan, serta pandangan kritis yang sangat bermanfaat dalam memperkaya dan mempertajam analisis dalam penelitian ini.
8. Seluruh Dosen Program Studi Informatika, yang telah membekali penulis dengan fondasi keilmuan dan berbagai pengalaman berharga selama masa perkuliahan.
9. Orang Tua dan Keluarga Tercinta, atas kasih sayang tak terhingga, doa yang tiada henti, serta dukungan moril maupun materiil yang menjadi sumber kekuatan utama bagi penulis dalam menyelesaikan studi ini.
10. Teman-teman seperjuangan “Banteng UwU” dan seluruh teman-teman Angkatan 2021, yang telah menjadi rekan diskusi, sumber semangat, dan tempat berbagi keluh kesah dalam suka maupun duka. Perjalanan ini terasa lebih ringan berkat kebersamaan kalian.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, segala bentuk kritik dan saran yang membangun akan diterima dengan senang hati demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi pengembangan ilmu pengetahuan, khususnya di bidang pemrosesan bahasa alami.

Surabaya, 22 Juli 2025

Penulis,

Rangga Widiasmara

NPM. 21081010085

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN	III
LEMBAR PERSETUJUAN	V
SURAT PERNYATAAN BEBAS PLAGIASI	VII
ABSTRAK	IX
ABSTRACT	XI
KATA PENGANTAR.....	XIII
DAFTAR ISI.....	XVI
DAFTAR GAMBAR.....	XX
DAFTAR TABEL	XXII
BAB I PENDAHULUAN.....	1
1.1. LATAR BELAKANG	1
1.2. RUMUSAN MASALAH	4
1.3. TUJUAN PENELITIAN	4
1.4. MANFAAT PENELITIAN.....	5
1.4.1. Manfaat Teoretis	5
1.4.2. Manfaat Praktis	5
1.4.3. Manfaat Lokal	6
1.5. BATASAN MASALAH	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. PENELITIAN TERDAHULU.....	7
2.2. EMOSI DALAM ANALISIS SENTIMEN	10
2.3. KLASIFIKASI EMOSI.....	10
2.4. EMOJI SEBAGAI FITUR DALAM ANALISIS SENTIMEN.....	11
2.5. NATURAL LANGUAGE PROCESSING (NLP).....	11
2.6. ANALISIS SENTIMEN.....	11

2.7. <i>CONVOLUTIONAL NEURAL NETWORK (CNN)</i>	12
2.8. LONG SHORT-TERM MEMORY (LSTM)	13
2.9. BIDIRECTIONAL LSTM (BiLSTM).....	15
2.10. CLASS WEIGHT	16
2.11 ARSITEKTUR HIBRIDA DEEP LEARNING	18
2.11.1 Hibrida Seri (Serial Hybrid).....	18
2.11.2 Hibrida Paralel (Parallel Hybrid)	19
2.12 SKEMA INPUT MODEL KLASIFIKASI	20
2.12.1 Model Input Tunggal (Single-Input Model)	20
2.12.2 Model Input Ganda (Multi-Input Model).....	20
BAB III DESAIN DAN IMPLEMENTASI SISTEM	21
3.1. METODE PENELITIAN	21
3.2. STUDI LITERATUR	22
3.3. PENGUMPULAN DATA	22
3.3.1. Akuisisi, Augmentasi, dan Penggabungan Data	23
3.3.2. Analisis Distribusi dan Skenario Pembagian Dataset	25
3.3.3. DATA STUDI KASUS KUALITATIF UNTUK PENGUJIAN SKENARIO REALISTIS	27
3.4. PRA-PEMROSESAN DATA	28
3.4.1. Pembersihan Data NaN dan Duplikat	28
3.4.2. Pemrosesan Emoji.....	29
3.4.3. Normalisasi dan Pembersihan Data Lanjutan (Teks).....	30
3.4.4. Tokenisasi	31
3.4.5. Encoding Label	31
3.5. PEMBANGUNAN MODEL	32
3.5.1 Model 1: Hybrid Seri Input Tunggal.....	33
3.5.2 Model 2: Hybrid Paralel Input Tunggal	34
3.6. PELATIHAN MODEL	37
3.6.1 Skenario Uji Coba	38
3.6.2. Proses Kompilasi dan Pelatihan Model	39

3.6.2.1 Kompilasi Model.....	39
3.6.2.2 <i>Callbacks</i>	40
3.6.2.3 Class Weighting	41
3.6.2.4 Batch Size.....	42
3.6.2.5 Pelaksanaan Proses Fit	42
3.6.3. Hyperparameter Tuning	43
3.6.3.1 Pentingnya Hyperparameter Tuning	43
3.6.3.2 Teknik Tuning.....	43
3.6.3.3 Hyperparameter yang Dapat Di-tuning.....	44
3.6.3.4 Proses Iteratif	44
3.7. EVALUASI MODEL.....	45
3.7.1. Dataset Evaluasi.....	45
3.7.2. Metrik Evaluasi	45
3.7.3. Analisis Hasil Evaluasi	46
BAB IV PENGUJIAN DAN ANALISA	49
4.1. LINGKUNGAN EKSPERIMEN	49
4.2. DESKRIPSI DATASET SETELAH PRA-PEMROSESAN	50
4.3. HASIL IMPLEMENTASI ARSITEKTUR MODEL	52
4.4. HASIL PELATIHAN MODEL DAN PEMBAHASAN	54
4.4.1. Analisis Kinerja Model 1 (Hybrid Seri Input Tunggal).....	55
4.4.2 Analisis Kinerja Model 2 (Hybrid Paralel Input Tunggal)	58
4.4.3. Analisis Kinerja Model 3 (Hybrid Seri Input Ganda).....	62
4.4.4. Analisis Kinerja Model 4 (Hybrid Paralel Input Ganda)	66
4.5. ANALISIS KOMPARATIF ANTAR MODEL	70
4.6. ANALISIS KUALITATIF KOMPARATIF DENGAN STUDI KASUS TERKONTROL.....	72
4.6.1 Uji Coba Skenario 1: Kalimat Normal.....	72
4.6.2 Uji Coba Skenario 2: Emoji Berbenturan	74
4.6.3 Uji Coba Skenario 3: Kalimat Sarkasme	76
4.6.4 Analisis Hasil Uji Coba Komparatif.....	78

4.7. DISKUSI TEMUAN	80
4.7.1. Efektivitas Pendekatan Multi-Input	80
4.7.2. Keunggulan Arsitektur Paralel.....	80
4.7.3. Analisis Anomali dan Tantangan yang Persisten.....	81
4.8. KETERBATASAN PENELITIAN	82
4.8.1. Keterbatasan Terkait Dataset	82
4.8.2. Keterbatasan pada Arsitektur dan Optimasi Model	83
4.8.3. Keterbatasan Lingkup Bahasa.....	84
4.8.4. Keterbatasan Sumber Daya	84
BAB V PENUTUP.....	85
5.1. KESIMPULAN	85
5.2. SARAN PENGEMBANGAN	86
DAFTAR PUSTAKA.....	89
LAMPIRAN.....	91

DAFTAR GAMBAR

GAMBAR 2.1 ARSITEKTUR CNN	13
GAMBAR 2.2 ARSITEKTUR LSTM.....	14
GAMBAR 2.3 ARSITEKTUR BILSTM	16
GAMBAR 3.1 TAHAP PENELITIAN.....	21
GAMBAR 3.2 PERSEBARAN KATEGORI EMOSI DALAM DATASET GABUNGAN	25
GAMBAR 3.3 ALUR PRA-PEMROSESAN DATA	28
GAMBAR 3.4 DIAGRAM KONSEPTUAL ARSITEKTUR MODEL 1	34
GAMBAR 3.5 DIAGRAM KONSEPTUAL ARSITEKTUR MODEL 2	35
GAMBAR 3.6 DIAGRAM KONSEPTUAL ARSITEKTUR MODEL 3	36
GAMBAR 3.7 DIAGRAM KONSEPTUAL ARSITEKTUR MODEL 4	37
GAMBAR 4.1: DISTRIBUSI JUMLAH SAMPEL PER KELAS EMOSI PADA DATASET FINAL	51
GAMBAR 4.2: KURVA AKURASI MODEL 1 SELAMA PELATIHAN	55
GAMBAR 4.3: KURVA LOSS MODEL 1 SELAMA PELATIHAN	56
GAMBAR 4.4: CONFUSION MATRIX KINERJA MODEL 1 PADA DATA UJI	58
GAMBAR 4.5: KURVA AKURASI MODEL 2 SELAMA PELATIHAN	59
GAMBAR 4.6: KURVA LOSS MODEL 2 SELAMA PELATIHAN	60
GAMBAR 4.7: CONFUSION MATRIX KINERJA MODEL 2 PADA DATA UJI	62
GAMBAR 4.8: KURVA AKURASI MODEL 3 SELAMA PELATIHAN	63
GAMBAR 4.9: KURVA LOSS MODEL 3 SELAMA PELATIHAN	64

GAMBAR 4.10: CONFUSION MATRIX KINERJA MODEL 3 PADA DATA UJI	66
GAMBAR 4.11: KURVA AKURASI MODEL 4 SELAMA PELATIHAN	67
GAMBAR 4.12: KURVA LOSS MODEL 4 SELAMA PELATIHAN	68
GAMBAR 4.13: CONFUSION MATRIX KINERJA MODEL 4 PADA DATA UJI	70

DAFTAR TABEL

TABEL 2.1 STUDI LITERATUR	7
TABEL 3.1 PEMBAGIAN EMOJI	24
TABEL 3.2 CONTOH DATASET	27
TABEL 4.1 KOMPOSISI PEMBAGIAN DATASET	50
TABEL 4.2 RINGKASAN PARAMETER	53
TABEL 4.3: LAPORAN KLASIFIKASI KINERJA MODEL 1 PADA DATA UJI	57
TABEL 4.4 LAPORAN KLASIFIKASI KINERJA MODEL 2 PADA DATA UJI	61
TABEL 4.5: LAPORAN KLASIFIKASI KINERJA MODEL 3 PADA DATA UJI	65
TABEL 4.6: LAPORAN KLASIFIKASI KINERJA MODEL 4 PADA DATA UJI	69
TABEL 4.7: PERBANDINGAN METRIK KINERJA UTAMA ANTAR KEEMPAT MODEL	70
TABEL 4.8 HASIL UJI COBA SKENARIO 1	73
TABEL 4.9 HASIL UJI COBA SKENARIO 2	75
TABEL 4.10 HASIL UJI COBA SKENARIO 3	77
TABEL 4.10 RINGKASAN AKURASI MODEL PADA SETIAP STUDI KASUS	78

Halaman ini sengaja dikosongkan